AMSAA SP-4

AD

AD 746245

ARMY MATERIEL SYSTEMS ANALYSIS AGENCY

COMPENDIUM ON RISK ANALYSIS TECHNIQUES

JULY 1972

Approved for public release;
distribution unlimited.

U S ARMY MATERIEL SYSTEMS ANALYSIS AGENCY
ABERDEEN PROVING GROUND, MARYLAND 21005

138

DISPOSITION

Destroy this report when no longer needed. Do not return it to the originator.

DISCLAIMER

The findings in this report are not to be construed as an official Department of the Army position.

WARNING

Information and data contained in this document are based on the input available at the time of preparation. The results may be subject to change and should not be construed as representing the AMC position unless so specified.

SPECIAL PUBLICATION NO. 4

COMPENDIUM ON RISK ANALYSIS TECHNIQUES

EDITED BY

Erwin M. Atzinger
Wilbert J. Brooks

AUTHORS

Erwin M. Atzinger
Wilbert J. Brooks
Michael R. Chernick
Brian Elsner
Ward V. Foster

July 1972

Approved for public release;
distribution unlimited.

RDT&E Project No. 1P765801MM1102

U.S. ARMY MATERIEL SYSTEMS ANALYSIS AGENCY
ABERDEEN PROVING GROUND, MARYLAND

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Risk analysis						
Decisions						
Systems analysis						
Program management						
Subjective probability						
Delphi						
Monte Carlo						
Network analysis						
Bayesian statistics						
Compendium						
Development program						
Materiel acquisition process						
Uncertainty						

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) U.S. Army Materiel Systems Analysis Agency Aberdeen Proving Ground, Maryland	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b. GROUP

3. REPORT TITLE

COMPENDIUM ON RISK ANALYSIS TECHNIQUES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Erwin M. Atzinger, Wilbert J. Brooks, Michael R. Chernick, Brian Elsner, and Ward V. Foster

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
July 1972	141	17

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. 1P765801MM1102	Special Publication No. 4
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release,; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY U.S. Army Materiel Command Washington, D.C.

13. ABSTRACT

The evolution of risk analysis in the materiel acquisition process is traced from the Secretary Packard memorandum to current AMC guidance. Risk analysis is defined and many of the existing techniques are described in light of this definition and their specific role in program management and systems analysis activities. Sections are included on Subjective Probability, Monte Carlo, Network Analysis, and Bayesian Statistics.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

Unclassified
Security Classification

U.S. ARMY MATERIEL SYSTEMS ANALYSIS AGENCY

SPECIAL PUBLICATION NO. 4

EMAtzinger/WJBrooks/MRChernick/
BElsner/WVFoster/sjj
Aberdeen Proving Ground, Md.
July 1972

COMPENDIUM ON RISK ANALYSIS TECHNIQUES

ABSTRACT

The evolution of risk analysis in the materiel acquisition process is traced from the Secretary Packard memorandum to current AMC guidance. Risk analysis is defined and many of the existing techniques are described in light of this definition and their specific role in program management and systems analysis activities. Sections are included on Subjective Probability, Monte Carlo, Network Analysis, and Bayesian Statistics.

 Next page is blank.

CONTENTS

CONTENTS (Continued)

Page

CHAPTER 2 (Continued)

CONTENTS (Continued)

Page

CHAPTER 4 (Continued)

CHAPTER 5 BAYESIAN STATISTICS

Next page is blank.

COMPENDIUM ON RISK ANALYSIS TECHNIQUES

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

In the past few years a great deal of attention has been focused on the Department of Defense's problems of unanticipated cost growth, system performance shortcomings and failure to deliver equipment on schedule -- all of which adversely reflect on the management, planning and systems analysis being done throughout the stages of materiel systems development and procurement. It is recognized within AMC that the first step in minimizing the occurrence of such problems is simply to acknowledge their potential existence. To this end, AMC is making an effort to consider program risks in decision making.

Formal recognition and emphasis on risk analysis in AMC resulted from a July 31, 1969 memorandum from Deputy Secretary of Defense, David Packard to the Secretaries of the Military Departments identifying problem areas in the weapon system acquisition process. In this memorandum, Secretary Packard cited inadequate identification and consideration of risks in major programs as a problem area requiring action. The initial formal recognition within AMC was part of the "Program for the Refinement of the Materiel Acquisition Process" (PROMAP-70) under the direction of the DCGMA, AMC.

While PROMAP-70 was instrumental in emphasizing this problem area to the AMC community, establishing a course of instruction in risk analysis, and involving the AMC community in pilot risk analyses, there still existed for some time a great deal of confusion concerning the role of risk analysis, who was responsible for conducting risk analyses, and what constituted an analysis of risk. Much of this confusion may have resulted because risk analysis was often promoted and perceived as a new concept instead of a forgotten or neglected part of systems analysis and/or program management. According to Quade and Boucher, risk analysis is explicitly defined as part of systems analysis. Their interpretation of systems analysis is as follows: "The idea of an analysis to provide advice is not new and, in concept, what needs to be done is simple and rather obvious. One strives to look at the entire problem, as a whole, in context, and to compare alternative choices in the light of their possible outcomes. Three sorts of inquiry are required, any of which can modify the others as work proceeds. There is a need, first of all, for a systematic investigation of the decision maker's objectives and of the relevant criteria for deciding among the alternatives that promise to achieve these objectives. Next, the alternatives need to be identified, examined for feasibility, and then compared in terms of their effectiveness and cost, taking time and risk into account. Finally, an attempt must be made to design better alternatives and select other goals if those previously examined are found wanting." (Reference 1) Consideration of risk in systems analysis is thought to be one of the major areas of application for risk analysis that Secretary Packard had in mind in his memorandum.

In addition to the preceding application of risk analysis in systems analysis, it is also thought that Secretary Packard was suggesting that risk analysis be applied by program managers in their daily activities. In this context, risk analysis need not have any broader objective than the identification and evaluation of program risks. In this case it is not necessary for it to be part of a systems analysis, but rather, it should be part of the management information that is routinely accumulated. This information should provide useful indicators to project managers about potential problem areas, and may be instrumental in program change decisions effected to reduce risk. It is thought that it is virtually impossible to realistically plan, schedule and control a development program without adequately identifying and evaluating risks on a daily basis. The risks, thus identified constitute the foundarion of information required to make major decisions.

The analysis of risk in program management and systems analysis at major decision points should be viewed as complementary activities in the successful development of a system. Neither activity by itself is sufficient for a successful development program, nor does doing both guarantee success. However, if risk analysis is conscientiously done in both areas, more informed decisions will result and the specific programs will be under tighter control. Unfortunately, these roles of risk analysis were not well defined at first, and instead of a renaissance of risk analysis in systems analysis and program management there ensued a concept definition period. During this period, a plethora of definitions surfaced, concept papers were written and there was a great deal of discussion about the subject. However, all of these efforts tended to cloud the concept more, and a general feeling of dissatisfaction resulted. Finally, the Commanding General, AMC tasked the Directorate for Plans and Analysis (AMCPA) to develop a set of guidelines for conducting risk analysis in the decision-making context which would distinguish this from risk analysis per se. The concept of decision risk analysis (DRA) has arisen from this effort. In addition, the role of risk analysis and decision risk analysis has been defined both in systems analysis and program management. This guidance, hopefully, places risk analysis in its proper context and should serve as the foundation for all future activities in this area.

---

[1] Quade, E.S., Boucher, W. I.; Systems Analysis and Policy Planning, Applications in Defense, American Elsevier Publishing Company, New York, 1968, p 11.

Decision risk analysis is defined as "a discipline of systems analysis, which in a structured manner, provides a meaningful measure of the risks associated with various alternatives," whereas risk analysis is defined as "an attempt to quantify uncertainty."

As the center for systems analysis activities within AMC, the U.S. Army Materiel Systems Analysis Agency (AMSAA) has worked very closely with AMCPA by participating in risk analyses, providing consultation to the AMC community, developing methods, and more recently preparing this compendium on risk analysis techniques. It is not our intent to develop a "cookbook" on risk analysis for we feel the diversity of problems makes this virtually impossible. Rather the objective here is to develop a definition of risk analysis in terms that allow one to relate to many existing risk analysis techniques.

1.2 DEFINITION OF RISK ANALYSIS

As mentioned previously, risk analysis can be conducted in either the day-to-day management of a program or as part of a systems analysis at some major decision point. In the latter context, risk analysis is one of many analyses -- such as resource and/or cost analysis and effectiveness which are usually associated with a systems analysis. These systems analysis -- in which the risks associated with various alternatives have been evaluated, are labeled decision risk analyses according to current AMC guidance.

On the other hand, risk analysis can be conducted to provide management information to the project office. The information obtained should be instrumental in planning, scheduling, and controlling the development program.

At this point, decision analysis as defined by Raiffa (Reference 2) will be differentiated from decision risk analysis in the AMC context. As defined previously, decision risk analysis is part of systems analysis. The term was originally coined by AMC to emphasize the risk analysis aspect of systems analysis. While decision risk analysis and systems analysis provide decision makers with a comprehensive and orderly presentation of choices for complex and unique problems, highlighting the uncertainty in the choice, they leave the decision maker the job of exercising his preferences in choosing an alternative. In complex problems, this is not an easy job. Decision analysis, on the other hand, elicits the decision maker's preferences (using utility theory) and incorporates these preferences in the selection of an alternative.

---

[2] Raiffa, Howard; Decision Analysis, Introductory Lectures on Choices Under Uncertainty, Addison-Wesley Publishing Company, Massachusetts, 1968.

While decision analysis is certainly a rational approach to decision making in private industry, the unique features of the materiel acquisition decision-making process make the concept difficult, if not impossible to apply. Determining a utility function for military decisions is thought to be more difficult because the most frequently used comparative measures are not as easy to relate to as profit and loss in the industrial context. In addition, it is also often very difficult to determine the identity of the decision maker. In most instances, there are actually a hierarchy of decision makers making choices at the various points in the decision-making process. Even if one chooses only to concentrate on the immediate decision maker he is generally unavailable to the analyst until the decision briefing. Thus, the decision analysis (utility theory) approach does not appear to be applicable in this process.

It is the authors' contention that the best one can do at this time is structure the decision risk analysis so that the trade-offs inherent in the alternatives are visible and meaningfully displayed.

Returning to the task at hand, recall that risk analysis was defined as an attempt to quantify uncertainty. While this definition is reasonable for general guidance, it is not sufficiently detailed to describe to the analyst what is involved in conducting a risk analysis. The intent here is to describe the sequence of activities involved in risk analysis.

In any risk analysis, there must be some variable(s) or area(s) of interest. Variables imply quantification. Areas, on the other hand, constitute the protions of a system or program which cannot be quantified. In the materiel acquisition process these variables or areas of interest fall into three broad categories; project completion times, project resource requirements/costs and systems effectiveness and performance. In the case where there exists a requirement or program objective for a variable or area of interest, risk is defined as the chance that the variable or area of interest will not meet the stated requirement or program objective. For example, in the development of a weapon system, one may be interested in the chance that the system will not be fielded by a planned date or within the budget or resource constraints.

In this context there is a sequence of three basic activities that comprise any risk analysis in either program management or systems analysis:

a. Identification of the variable(s) or area(s) of interest;

b. Consolidation of all information about the variables(s) or area(s) of interest; and

c. Evaluation of the risks.

The identification of the variable(s) or area(s) of interest is certainly one of the most important activities in any risk analysis. If all of the important variables or areas of interest have not been identified, the risk evaluation may not assess the dominant risks. In a quantitative evaluation, this means that the risks have been underestimated, and in a qualitative assessment this means that all potential problems may not have been considered. Unfortunately, very little guidance can be given to the analyst on the identification activity. It can only be recommended that the analyst examine the history of similar programs, become intimately familiar with the system in question and systematically and thoroughly investigate all aspects of the current program. The following is a suggested list of items which should be examined at a minimum:

a. Requirements document (MN),

b. Program budget,

c. Program plan,

d. Design specifications,

e. System history,

f. Current problems, and

g. Similar systems.

In general, the primary sources of these data will be the developing agency, testing agency, and the primary contractors.

Note that there are certain cases in which this identification stage may seem unnecessary. For instance, if the directive is to evaluate the risk associated with a specific set of variables, no identification seems required. In such a case, however, the analyst must be careful to adequately address all the interdependences that may exist among variables. In order to assess the risk associated with a specific variable, he may be forced to become intimately familiar with the entire program. Hence, the identification activity is still required.

Having identified the variables(s) or area(s) of interest, the next steps are to consolidate and evaluate all information concerning these factors. Depending on the particular application, this consolidation and evaluation may be either quantitative or qualitative.

In certain applications, quantitative measures may be meaningless. For example, in system design, the developing agency will be interest in pinpointing the potential problem areas of a particular design. The consequences of these problems will then be evaluated so that fallback positions might be developed and tighter management controls can be implemented for these high risk activities. For example, consider the development of a component of a tank for which the design is well within the state of the art. Assume, however, that the contractor is planning to design and implement a new automated production line for this component. Any problems encountered in the development of automated production line will have a direct impact on the development of the component. What needs to be monitored quite closely is not the development of the component, but the development of the new production line. The identification of this potential problem area and the evaluation of the possible consequences of problems in developing this new production line certainly constitutes risk analysis even though the chance of not meeting a program requirement or objective is not addressed explicitly or necessarily quantitatively.

On the other hand, many instances will occur for which consolidation and evaluation can be handled quantitatively. The variable of interest may be a random quantity or a fixed unknown constant which must be estimated by a random quantity. In either case, the primary objective of the consolidation phase is to obtain (objectively or subjectively) a representative probability distribution reflecting the uncertainty in the variable. Having obtained this probability distribution, the analyst is in a position to meaningfully evaluate the risk or the chance that the variable will not meet the stated requirement or program objective.

The next section includes a summary of the specific techniques addressed in this compendium as they relate to previously defined sequences of activities of risk analysis.

1.3 CONTENTS OF THE COMPENDIUM

There are four sections in the compendium. They are

a. Subjective Probability,

b. Monte Carlo,

c. Network Analysis, and

d. Bayesian Statistics.

Each of these sections has been written such that the comprehension of any section depends only on the information contained it it and in the Introduction.

On the other hand, the reader, who is interested in all of these sections, is encouraged to read the sections in their order of appearance. Although there is no rigid format for each section, the following information is included:

a. General introduction to the technique(s),

b. Description of the technique(s),

c. Discussion of the relationship of the technique(s) to the three risk analysis activities,

d. Discussion of the pros and cons of the technique(s), and

e. Summary and conclusions.

In the next few paragraphs, each of the preceding techniques will be described relative to the three general risk analysis activities and the situations that frequently confront the analyst, especially in quantitative consolidation.

In many risk analyses few data are available to estimate the distribution of the variable of interest. In these instances subjective techniques can be meaningfully employed during the consolidation activity to quantify information in the form of a probability density function.

Several of these techniques are categorized by the assumptions about the expert's level of understanding of basic probability concepts and by the unique questioning procedure used to elicit the expert's judgement.

Of course there may exist situations where a group of experts is available. In this same subjective probability section, a modified Delphi procedure is presented as a method for consolidating group judgment when estimating an unknown probability distribution.

The Delphi procedure may also be applicable in other identification activities and in situations where comparison of alternatives is difficult because quantitative models are inadequate. Although not risk analysis activities, these type applications present an interesting method for answering important questions in an uncertain environment.

A Monte Carlo procedure section is also included in the compendium. Historically, Monte Carlo procedures have been invaluable in situations where quantification of the uncertainty in a variable is complicated by the fact that the variable is itself a non-trivial function of several other variables. This situation commonly occurs in the consolidation activity of many performance risk analyses.

Monte Carlo procedures are also used extensively in network analysis to simulate network representations. These network analysis techniques are an integral part of the consolidation activity in risk analysis and allow one to model not only time and cost uncertainty, but uncertainty in future events as well. Network analysis is particularly useful in evaluating time and cost risks in development programs. Two network analyzer programs, PERT and RISCA, are described in detail in the compendium. They were selected as being representative of the techniques for handling different types of network representations for different purposes.

While PERT analyses are generally used in project management as a tool i planning, scheduling, and controlling program activities, RISCA provides a framework for modeling project schedule, cost, and event uncertainties for specific decision-making purposes. In the context of risk analysis, RISCA provides a method for quantifying, in meaningful summarized form, development time and cost risks. In addition, network analysis techniques, like RISCA, can provide the foundation for structuring a decision risk analysis.

The final section is devoted to Bayesian Statistics. Bayesian Statistics enjoys a unique position in risk analysis. There frequently exist situations where the analyst has both data and expert judgment to draw upon in constructing the probability distribution of interest in the consolidation activity. Bayesian statistics provides the analyst with a tool for synthesizing all of this information into one probability distribution which can then be used to directly estimate risks.

The authors recognize that many people have given considerable thought to risk analysis in the materiel acquisition process. Only recently, however, has the concept of risk analysis been placed in perspective in the activities of program management and systems analysis. The definition of risk analysis presented in this compendium is an attempt to bridge the gap between general guidance and practical application.

CHAPTER 2

SUBJECTIVE PROBABILITY

2.1 INTRODUCTION

"The subjective or personalistic concept of probability is relatively recent.* Its application to statistical problems has occurred virtually entirely in the post World War II period, particularly in connection with statistical decision theory. According to this concept, the probability of an event is the degree of belief or degree of confidence placed in the occurrence of an event by a particular individual based on the evidence available to him. This evidence may consist of relative frequency of occurrence data and any other quantitative or non-quantitative information. If the individual believes it is unlikely an event will occur, he assigns a probability close to zero to its occurrence; if he believes it is very likely the event will occur, he assigns it a probability close to one.

"Those who accept subjective probability argue that in assigning probabilities to events, other information in addition to past relative frequencies of occurrence should be taken into account. To make this point clear, let us consider an oversimplified, somewhat artificial example. Suppose a company which purchases a product from a certain supplier has had the following experience with shipments from that firm: 1 percent defective items in each of 10 shipments, 2 percent defective in each of 85 shipments, and 3 percent defective in each of 5 shipments. Assume all shipment contained the same number of items. These data are displayed in Table 2.1

TABLE 2.1 PERCENTAGE OF DEFECTIVE ITEMS IN ONE HUNDRED SHIPMENTS

Percent Defectives	Number of Shipments
1	10
2	85
3	5
TOTAL	100

*The concept was first introduced in 1926 by Frank Ramsey who presented a formal theory of personal probability in F. P. Ramsey, The Foundation of Mathematics and Other Logical Essays (London: Kegan Paul; New York: Harcourt, Brace, & World, 1931). The theory was developed primarily by de Finetti, Koopman, I. J. Good, and L. J. Savage.

"Suppose the purchasing company wants to know the probability that the next shipment from this supplier will contain 2 percent defective items. In the absence of any further information, it seems reasonable to assign a probability of 0.85 to that event. That is, since shipments with 2 percent defectives occurred in 85 percent of the past cases, the relative frequency of occurrence would seem to be a good estimate of the probability in question. However, suppose the purchasing company acquires some additional information. It learns that the engineer who has been in charge of production for the supplier, and who has been the key person responsible for the maintenance of the quality level of the product has just resigned his position with the company. Furthermore, it is known that his knowledge has not been passed on to a suitable replacement. Therefore, a deterioration in quality of the product, at least until suitable remedial measures can be instituted, seems reasonable. Should a probability of 0.85 still be assigned?

"In this case it certainly seems reasonable that the assignment of probabilities should no longer depend solely on past relative frequency data. The purchaser, as a practical business man, should undoubtedly anticipate that shipments in the near future will display quality levels different from those indicated by the data in Table 2.1. For example, percentages of defectives in excess of 3 percent are possibilities for shipments in the near future, and somehow or other, for decision-making purposes, the purchaser must reckon with the likelihood of such shipments. What the purchaser now needs is a new distribution of all the percentage defectives he feels are possible with probability assignments attached to each. It might be argued that the purchaser should wait until conditions within the supplier company are again stable and reasonable assurance is given that acceptable quality levels will be maintained. However, suppose the purchaser cannot delay his decisions for that period, and must take appropriate action now.

"Subjective probabilities should be assigned now on the basis of all objective and subjective evidence currently available. These probabilities should reflect the decision maker's current degree of belief. Reasonable persons might arrive at different probability assessments because of differences in experience, attitudes, values, etc. Furthermore, in general, these probability assignments may be made for events which will occur only once, in situations where the concept of a repetitive sequence of trials under uniform conditions does not appear to be a useful model.

"This approach is thus a very broad and flexible one, permitting probability assignments to events for which there may be no objective data, or for which there may be a combination of objective and subjective data. These events may occur only once and may lie entirely in the future. However, the assignments of these probabilities must be consistent. For example, if the purchaser in the illustration above assigns a probability of 0.40 to the event that a shipment will have 2 percent or less defective items, then a probability of 0.60 must be assigned to the event that a shipment will have more than 2 percent defective items." (See Reference 1).

The following section describes how the subjective probability concept may be applied to the consolidation activities in a risk analysis.

In conducting a risk analysis either to generate management information or as a part of a systems analysis, the consolidation activities may be complicated by a complete lack of data or the existence of very little data from which to estimate the variable(s). The variables of interest may be random variables or unknown quantities, but in either case, it must be assumed that the variables have been adequately identified. In the case of the unknown quantity, the estimate of the quantity is a random variable. In situations where no data or very few data exist, subjective probability provides the only alternative to the analyst in his effort to quantify the uncertainty in these variables. The personnel from whom such personalistic probabilities are elicited are usually engineers and scientists associated with the development of specific components and subsystems of the overall weapon system. An overview of some methods of eliciting subjective probability estimates and subsequent distributions from these experts is presented in Section 2.2.

Throughout the remainder of this chapter, the discussion of methods and examples will be in terms of performance variables; however, the reader should note that these methods will apply analogously when evaluating time and cost variables. Performance characteristics are analyzed here only for purposes of continuity and comparability.

At the total system level, the appropriate variables representing overall system performance can usually be expressed in terms of a few critical system performance characteristics. These characteristics indicate the system's capability in performing certain predefined missions. Aircraft performance characteristics, for example, might include speed, range, and altitude. However, in order to achieve estimates for the performance characteristics at the system level, it is first necessary to derive functional relationships for these system performance characteristics (i.e., the dependent variables) in terms of the appropriate subsystem component characteristics (i.e., the independent variables).

The next stage is to interrogate the most technically qualified people involved in each area of the system development regarding their appraisal of component characteristic variability. Because objective data are generally sparse in a development program, these

[1]Hamburg, Morris; Statistical Analysis for Decision Making, Harcourt, Brace & World, Inc., New York, 1970, pp 12-13.

component characteristic appraisals will be in the form of subjective probability distributions which will provide the basis for estimating system performance characteristic uncertainty.

If possible, the subjective assessment should be made at the subsystem component level where engineering experts will find it easiest to relate to characteristic uncertainty. Simulation procedures (e.g., Monte Carlo) can then be applied using the appropriate functional relationships to obtain a measure of the uncertainty involved at the system level.

For the purpose of bringing the derivation of subjective probability estimates into perspective, consider the following example. Suppose the development program involves a certain type aircraft where the performance characteristics critical to the aircraft's mission capability are defined to be speed, altitude, range, and endurance. These performance variables which reflect overall system performance are listed across the top of Table 2.2.

For each of these performance characteristics a design equation or relationship must be developed which reflects the influence of the subsystem component characteristics at the system level. Table 2.2 describes which component characteristics (W_g, W_f,...etc.) are involved in the design equations for each of the project performance characteristics in the aircraft example.

The relationship for V_{max} (the maximum, constant altitude, level flight speed of an aircraft), for example, may be of the form:

$$V_{max} = \left\{ \frac{195.5 \left[T_{max} + \left(T_{max}^2 - 1.274 \dfrac{C_{D_o} S W_g^2}{oeb^2} \right)^{1/2} \right]}{S C_{D_o}} \right\}^{1/2}$$

For each of the component characteristics in this equation, subjective estimates of its probability distribution must now be obtained. For the aircraft example, Table 2.3 depicts a hypothetical set of distributions for the relevant components defined in Table 2.2. Simulation procedures (e.g., Monte Carlo) can now be applied using the appropriate functional relationship to obtain the distribution of the performance characteristic at the system level.

The following section provides a brief overview of several of the more common techniques for eliciting subjective probabilities at the component level. Each of these will then in turn be discussed in detail.

TABLE 2.2 PERFORMANCE AND COMPONENT CHARACTERISTICS FOR HYPOTHETICAL AIRCRAFT DEVELOPMENT PROJECT[2]

		Performance Characteristics			
Component Characteristics	Units	Speed (V_{max}) (mi/hr)	Altitude[a] (H_{max}) (ft)	Range (R_{max}) (mi)	Endurance (E_{max}) (hr)
W_g	lb	x	x	x	x
W_f	lb			x	x
C_{D_o}			x		
b	ft	x	x	x	x
S	ft^2	x	x	x	x
e		x	x	x	x
T	lb	x	x		
α			x		
C'	lb/hr/lb of T			x	x

W_g - initial gross weight

W_f - final gross weight

C_{D_o} - drag coefficient for zero lift

b - wing span

S - wing area

e - efficiency factor

T - thrust

α - thrust lapse rate factor

C' - fuel consumption

[2]Timson, F. J.; Measurement of Technical Performance in Weapon System Development Programs: A Subjective Probability Approach, Memorandum RM-5207-ARPA, December 1968, The Rand Corporation, p 15.

TABLE 2.3 PROBABILITY DISTRIBUTIONS FOR COMPONENT CHARACTERISTICS (INPUTS) IN HYPOTHETICAL AIRCRAFT DEVELOPMENT PROJECT[2]

W_g (lb)	250,000	260,000	270,000	280,000	290,000
W_f (lb)	130,000	140,000	150,000	160,000	170,000
p(W)	.05	.10	.35	.30	.20
C_{D_o}	.015	.016	.017	.018	--
$p(C_{D_o})$	.10	.40	.30	.20	--
b (ft)	170	175	180	185	190
S (ft^2)	3,500	3,600	3,700	3,800	3,900
p(S)	.15	.15	.35	.25	.10
e	.70	.75	.80	.85	.90
p(e)	.20	.25	.25	.15	.05
T (lb)	65,000	70,000	75,000	80,000	--
p(T)	.15	.35	.25	.25	--
α	1.2	1.4	1.6	1.8	--
p(α)	.20	.25	.30	.25	--
C' (lb/hr/lb of T)	.80	.85	.90	.95	--
p(C')	.30	.30	.25	.15	--

[2] *Ibid.*, p 16.

2.2 OVERVIEW OF THE TECHNIQUES

2.2.1 Choice-Between-Gambles Technique for Deriving Probability Density Functions.

This method employs betting-type or gambling situations to elicit inferred probability of occurrence responses from the expert. The expert proceeds to reveal indifference probabilities between a hypothetical gamble and a real-life gamble involving a fixed level of the variable of interest. By varying probabilities in the hypothetical gamble and the level of the variable of interest, a subjective probability distribution is obtained.

2.2.2 Choice-Between-Gambles Technique for Deriving Cumulative Distribution Functions.

A cumulative distribution function of subjective probabilities is derived based on the expert's revealed indifference characteristic values. These values result from a hypothetical gamble versus real-world-gamble (i.e., involving the variable of interest) betting situation for a fixed level of probability. Each successive decision stage of the procedure reveals a characteristic value within a specified interval of values which divides the interval into equally probable sub-intervals. Relating each specified value directly to a cumulative probability of occurrence, a distribution function is obtained.

2.2.3 Standard Lottery.

A probability density function is derived for the component characteristic variable of interest. Probabilities are inferred based on a selected number of hypothetical lottery tickets chosen from a lot of fixed size. The number of tickets chosen by the expert for each defined level of the component characteristic directly infers his subjective feeling for the probability of realization of that characteristic value.

2.2.4 Modified Churchman-Ackoff Technique.

No indifference assessments or betting decisions are required in this technique. Instead, the expert is asked to make relative probability-of-occurrence-type judgments (i.e., greater than, equal to, and less than) between various sets of possible characteristic probabilities. Then, he is asked to make numerical relative probability judgments between values on the ordinal scale desired in the previous decision stage. The resulting relative probability scale is directly converted algebraically into a probability density function. This technique is applicable only if the expert has an understanding of probability theory.

2.2.5 Modified Delphi Technique.

Group (i.e., at least 3 experts) subjective probability distributions, as opposed to individual probability distributions, are desired. Employing the Modified Delphi Technique, individual probability responses are elicited, reasons stated regarding such judgements are made, and all information is fed back to all respondents in an iterative procedure. A group probability response for all characteristic values is ultimately defined by averaging.

The techniques developed in this section for eliciting subjective probabilities involve either asking the expert

(1) to make choices between different betting situations,

(2) state preferences between combinations of component characteristic values or

(3) evaluate responses in a group decision-making situation.

The resulting probability distributions are in the form of either a probability density function or a cumulative distribution function. Since each type of function is derivable from the other, the determining criterion for selecting a given technique is not the form of the output but the expert's relative ability to respond meaningfully and validly to each type of questioning and the expert's familiarity with the method. In addition, it is imperative that the analyst administering the interrogation be a skilled interviewer with an ability to establish the necessary rapport required to promote candid discussions with the appraisers.

A detailed discussion of each of the five questioning techniques will now be presented.

2.3 CHOICE-BETWEEN-GAMBLES TECHNIQUE FOR DERIVING PROBABILITY DENSITY FUNCTIONS

2.3.1 Introduction.

The objective of the Choice-Between-Gambles Technique is to subjectively derive a discrete probability density function of component characteristic achievement levels (i.e., the probability that a component characteristic will achieve a specified level). The inputs for eliciting these probability responses from the experts are composed of choice-between-gambles or betting-type questions administered by the analyst. It is believed by many authors in the field of subjective decision making that this form of questioning results in a more realistic subjective density function than a direct questioning approach. This latter method of asking the expert directly what probability he attaches to a particular outcome for a component characteristic, although simple in application, has little likelihood of success in most cases. Many individuals either have no ability to think directly in terms of probabilities, or have difficulty communicating them without the aid of an auxiliary tool such as the one employed in this technique.

2.3.2 Description.

The technique is an iterative procedure which is initiated by presenting two alternative gambling situations. The expert is asked to choose between a real-world gamble involving values of a component characteristic of the project in question with unspecified probabilities and a hypothetical gamble involving two objective events, E_1 and E_2 with given objective probabilities, $P(E_1)$ and $P(E_2)$. The monetary payoffs for both gambles are made equal to facilitate the expert's ability to discriminate. Next, the probabilities of the hypothetical gamble are varied (starting with equal probabilities for E_1 and E_2) until the expert is indifferent between the two gambles. Hence, the appraiser's subjective probabilities regarding the outcomes of the real-world gamble are inferred by the resulting probabilities from the hypothetical gamble.

As an illustration of the above technique, consider a real-world gambling situation involving two possible thrusts of a jet engine under development and a hypothetical gamble with possible events E_1 and E_2. The payoffs are stated as: (1) \$10 if one thrust is realized, and \$0 if the other thrust occurs; and (2) \$10 if event E_1 occurs, and \$0 if event E_2 is realized. Table 2.4 reflects this initial decision situation.

TABLE 2.4 DECISION SITUATION

Real-World Gamble			Hypothetical Gamble		
Payoff	\$10	0	Payoff	\$10	0
Thrust	36,000 lb $\pm$1,000 lb	not 36,000 lb $\pm$1,000 lb	Event	E_1	E_2
Probabilities	?	?	Probabilities	0.5	0.5

The assumption is that if the expert chooses the real-world gamble, he will receive \$10 if a thrust of 36,000 $\pm$ 1,000 lb actually occurs, and \$0 if any other thrust is realized. If he selects the hypothetical gamble, he will receive \$10 if E_1 occurs and \$0 if E_2

occurs. Therefore, if his decision in the first round is to accept the real-world gamble, then it is immediately inferred that his subjective probability assessment that a thrust of 36,000 $\pm$1,000 lb will be achieved is greater than 0.5. Thus, in the next decision rounds the analyst will adjust the probability of occurrence of hypothetical event E_1 upward, and that for event E_2 downward. This procedure is then continued in an iterative fashion to the stage where the expert is indifferent to the two gambling situations. Suppose that this stage is ultimately achieved at $P(E_1) = 0.7$, $P(E_2) = 0.3$, then P(thrust = 36,000 $\pm$1,000) is inferred to be 0.7. This revised decision situation is depicted in Table 2.5.

TABLE 2.5 REVISED DECISION SITUATION

Real-World Gamble			Hypothetical Gamble		
Consequence	\$10		Consequence	\$10	0
Thrust	36,000 lb $\pm$1,000 lb	not 36,000 lb $\pm$1,000 lb	Event	E_1	E_2
Probabilities	?	?	Probabilities	0.7	0.3

Having obtained the probability of thrust equal to 36,000 $\pm$1,000 lb, the next step is to change the thrust in the real-world gamble to the next interval that the expert will be able to discriminate between its probability of occurring over that of the previous value. At each successive stage, then, probabilities are derived for various interval values of thrust. Finally, each endpoint of the density function is determined when the expert is indifferent between the two gambles, with $P(E_1) = 0$, and $P(E_2) = 1$.

The resulting probability distribution for this example could be as shown in Table 2.6.

TABLE 2.6 FINAL PROBABILITY DISTRIBUTION

Thrust (lb)	Probability
32,000 $\pm$1,000	0.0
34,000 $\pm$1,000	0.2
36,000 $\pm$1,000	0.7
38,000 $\pm$1,000	0.2
40,000 $\pm$1,000	0.0

has little likelihood of success in most cases. Many individuals either have no ability to think directly in terms of probabilities, or have difficulty communicating them without the aid of an auxiliary tool such as the one employed in this technique.

2.3.2 Description.

The technique is an iterative procedure which is initiated by presenting two alternative gambling situations. The expert is asked to choose between a real-world gamble involving values of a component characteristic of the project in question with unspecified probabilities and a hypothetical gamble involving two objective events, E_1 and E_2 with given objective probabilities, $P(E_1)$ and $P(E_2)$. The monetary payoffs for both gambles are made equal to facilitate the expert's ability to discriminate. Next, the probabilities of the hypothetical gamble are varied (starting with equal probabilities for E_1 and E_2) until the expert is indifferent between the two gambles. Hence, the appraiser's subjective probabilities regarding the outcomes of the real-world gamble are inferred by the resulting probabilities from the hypothetical gamble.

As an illustration of the above technique, consider a real-world gambling situation involving two possible thrusts of a jet engine under development and a hypothetical gamble with possible events E_1 and E_2. The payoffs are stated as: (1) \$10 if one thrust is realized, and \$0 if the other thrust occurs; and (2) \$10 if event E_1 occurs, and \$0 if event E_2 is realized. Table 2.4 reflects this initial decision situation.

TABLE 2.4 DECISION SITUATION

Real-World Gamble			Hypothetical Gamble		
Payoff	\$10	0	Payoff	\$10	0
Thrust	36,000 lb $\pm$1,000 lb	not 36,000 lb $\pm$1,000 lb	Event	E_1	E_2
Probabilities	?	?	Probabilities	0.5	0.5

The assumption is that if the expert chooses the real-world gamble, he will receive \$10 if a thrust of 36,000 $\pm$ 1,000 lb actually occurs, and \$0 if any other thrust is realized. If he selects the hypothetical gamble, he will receive \$10 if E_1 occurs and \$0 if E_2

The total of the preceding subjective probabilities equals 1.1 -- a result that obviously does not adhere to the probability axiom which states that the sum of the weights assigned to any set of mutually exclusive and collectively exhaustive events will equal 1.0. Thus, to resolve this conflict, the analyst can either: (1) reassess the expert's probabilities, or (2) normalize the derived probabilities by dividing each one by the sum of all the subjective probabilities.

The potential advantages, limitations, and assumptions pertaining to the Choice-Between-Gambles Technique are now presented.

2.3.4 Advantages.

This gambling approach derives probability density functions through inference rather than direct questioning. As noted earlier in the discussion, it appears that such an organized step-wise procedure for eliciting judgments, which allows the engineer or scientist to choose between alternatives rather than make direct probability judgments, results in a more valid density function.

Compared to most other techniques discussed in this section, this technique is not time consuming in its application. It is simple to apply and results directly in a probability density function.

2.3.5 Limitations.

As with the resulting distributions from all other techniques discussed in this section, this technique produces only discrete probability distributions. Of course, it is possible to obtain continuous subjective probability distributions but the techniques require that the subject be able to evaluate the entire distribution at one time. This is not as easy as dealing with the distribution in pieces. Thus, continuous distributions are not discussed.

In addition, the expert may find it difficult to determine the highest or lowest characteristic value for which he can state a subjective probability due to his limited ability to discriminate between values. However, a simple procedure for determining successive discrete values is given below:

a. Start with the preceding value that has been given a probability assessment.

b. Progress upward (or downward) on the scale of values until the expert is able to state a simple probability preference (greater than or less than) regarding the relative probabilities of occurrence of the two characteristic values. If he is thus able to state such a preference, then it is inferred that he is able to state such a preference, then it is inferred that he is able to discriminate between the two values.

c. Employ the Choice-Between-Gambles Technique on this new value.

d. Return to step a if the probability derived in c is greater than 0; otherwise, stop at the last iteration.

As for all the techniques in this section, the resulting probabilities represent the engineer's beliefs at a particular time under existing conditions.

2.3.6 Assumptions.

It is assumed the monetary rewards offered as consequences for correct responses are sufficient in magnitude to motivate the appraiser in forming his judgments. It is assumed also that the expert's concern for the project success, his integrity, and his decision-making abilities contribute also to the degree to which his judgments represent his personal beliefs.

It is also assumed that the appraiser is knowledgeable and experienced in his field, and is sufficiently well-founded in probability theory to respond meaningfully to the questioning procedure.

2.4 CHOICE-BETWEEN-GAMBLES TECHNIQUE FOR DERIVING CUMULATIVE DISTRIBUTION FUNCTIONS

2.4.1 Introduction.

The objective of this technique is to subjectively derive a cumulative distribution function reflecting the chance of occurrence of a range of component characteristic values up to and including a specified limit value. Instead of changing the probabilities of occurrence of hypothetical events to arrive at indifference between two gambles (as in the previous technique), this technique involves fixed probabilities of occurrence of hypothetical events, while the component characteristic values of the real-world gamble are changed until indifference is achieved.

2.4.2 Description.

Figure 2.1 shows that the first decision stage of this procedure involves: (1) a hypothetical gamble with events E_1 and E_2 having equal probabilities of occurrence (i.e., $P(E_1) = 0.5$, and (2) a "real world" gamble with two bets b_1 (the bet that the actual component characteristic will realize a value greater than T_1).

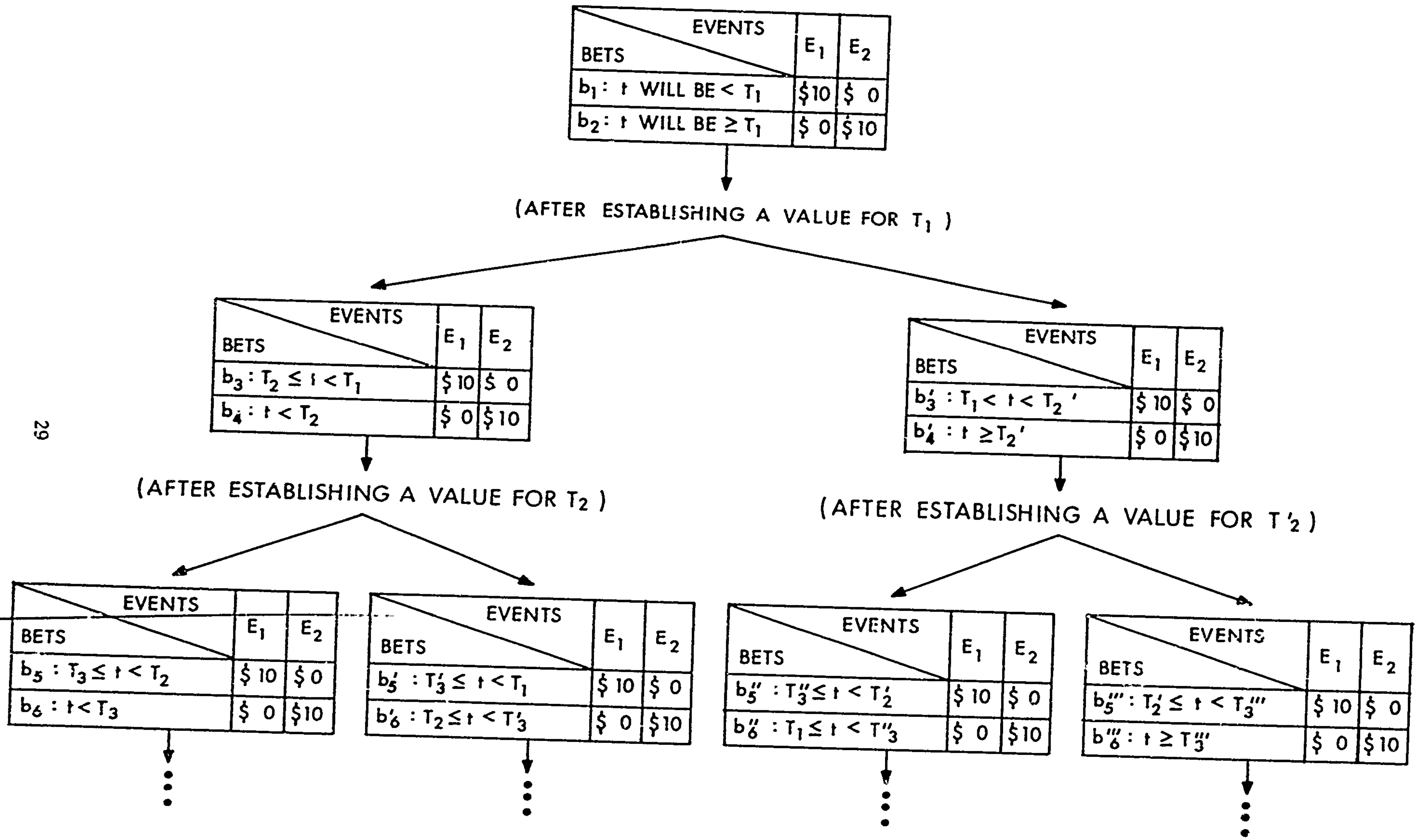


Figure 2.1. Series of Choices Between Bets to Determine a Probability Distribution Function[2]

2. Timson, Op. Cit., p 51.

As in the previous technique, the objective at each stage is to achieve indifference between the two gambles presented. To achieve this, the actual component characteristic value T_1 (related to bets b_1 and b_2) is varied until the expert cannot state a preference between the real gamble and the hypothetical gamble. Thus, when indifference occurs at the first stage, it is automatically inferred that the probability the component characteristic value will be less than T_1 is 0.5 and the probability that the value will be greater than or equal to T_1 is 0.5. This condition implies that the expert has a 50-50 chance of earning the \$10 outcome and a 50-50 chance of earning \$0 if he chooses b_1 or b_2.

Having established T_1 (the value which divides the density function in half) the next lower stage (See Figure 2..) involves two gambling situations. The objective here is to divide the two halves of the density distribution (as defined in stage 1) into two more halves. This is accomplished by deriving T_2 (the characteristic value which represents the 0.25 probability level of the function) and T_2' (which represents the 0.75 probability level). Thus, upon completion of decision stage 2, the values T_2, T_1, and T_2', divide the probability density function into fourths with the probability that the value will be less than T_2 is one-fourth, the probability the value will be greater than or equal to T_2 and less than T_1 is one-fourth, the probability it will be greater than or equal to T_1 and less than T_2' is one-fourth, and the probability it will be greater than or equal to T_2' is one-fourth.

As Figure 2 1 indicates, the equal portions of the density function are successively halfed at each subsequent stage of decision making until the expert is no longer able to discriminate between values.

The hypothetical values presented in Table 2.7 (as a continuation of the engine thrust example of the previous section) give values resulting from the three successive stages of questioning.

TABLE 2.7 HYPOTHETICAL RESULTS OF QUESTIONING PROCEDURE ILLUSTRATED IN FIGURE 2.1[2]

T_i	Value (lb)
T_1	38,000
T_2	36,000
T_2'	40,000

[2] Timson, *Op. Cit.*, p 52.

TABLE 2.7 (Continued)

T_i	Value (lb)
T_3	35,500
T_3'	37,000
$T_{3''}$	39,500
$T_{3'''}$	40,500

The correscponding probability distribution is displayed in Table 2.8.

TABLE 2.8 PROBABILITY DISTRIBUTION FUNCTION[2]

T_i	$P(t \leq T_i)$
35,500	0.125
36,000	0.250
37,000	0.375
38,000	0.500
39,500	0.625
40,000	0.750
40,500	0.875

The last step in the procedure is the derivation of the endpoint values of the distribution. The lower end-point of the distribution (T_L) will satisfy the condition that the probability of the actual value of thrust being greater than T_L is 1.0. The upper end-point (T_U) will represent the lowest value of thrust for which the probability of the actual value being less than T_U is 1.0. The technique for deriving a density function as outlined in the last section is useful for determining these endpoints. For the thrust example, the appropriate values might appear as in Tables 2.9 and 2.10.

[2]Timson, Op. Cit., p 52.

TABLE 2.9 LOWER LIMIT

Real-World Gamble			Hypothetical Gamble		
Consequence	\$10	\$0	Consequence	\$10	\$0
Thrust	≤ 35,000 lb	anything else	Event	X	Y
Probability	?	?	Probability	0.0	1.0

TABLE 2.10 UPPER LIMIT

Real-World Gamble			Hypothetical Gamble		
Consequence	\$10	\$0	Consequence	\$10	\$0
Thrust	≥ 42,000 lb	anything else	Event	X	Y
Probability	?	?	Probability	0.0	1.0

From these values, the probability distribution function in Table 2.11 is inferred directly.

TABLE 2.11 PROBABILITY DISTRIBUTION FUNCTION OF TABLE 2.7 WITH END-POINTS[2]

T	$P(t \leq T)$
35,000	0.000
35,500	0.125
36,000	0.250
37,000	0.375
38,000	0.500
39,500	0.625
40,000	0.750
40,500	0.875
42,000	1.000

[2] *Loc. Cit.*, p 53.

2.4.3 Advantages.

The cumulative distribution function derived by this technique can serve the same purposes as the probability density function in the previous section due to the fact that one function can be derived from the other. Therefore, when both questioning procedures are employed on the same individuals, both techniques together can provide a check on the consistency of the expert's responses.

This technique also offers an alternative gambling approach to the one previously described. Instead of varying the probabilities of occurrence, a characteristic value within a given interval of values is varied until the expert is indifferent between the occurrence of the two groups of sub-intervals on each side of the specified value with respect to fixed probabilities. Thus, for certain appraisers, it may be easier to think in terms of the areas under various portions of the density function rather than in terms of the probability points on the density curve. It is to be understood that the expert is not required to actually visualize the entire density curve. However, in the course of the questioning procedure, he will in effect be revealing equal probability intervals.

The probability considered in each betting situation remains fixed at 0.50. This condition is the easiest for an individual to comprehend because of common encounters in real-life with random processes such as coin-tossing.

2.4.4 Limitations.

The primary limitation of this technique lies in the expert's ability to respond when through further sub-division the probability of occurrence of the interval becomes small.

In addition, determining the end point (0.0 and 1.0 cumulative probability values) of the distribution requires the use of other techniques mentioned in this section.

Adjustment of the final distribution table of values must be made before sampling from the distribution. Looking at Table 2.10, the probability that the thrust will be between 35,500 lb and 36,000 lb is 0.125. Therefore, for any randomly sampled probability between 0.125 and 0.250, the thrust thus chosen should be between 35,500 and 36,000 lb. A logical choice would be the midpoint, 35,750 lb. This approach provides a reasonable solution to the problem.

2.4.5 Assumptions.

The following assumptions are made whenever the technique is applied:

1. The expert will be able to make more rational, consistent, and correct judgments when presented with betting situations than if he were asked to state probabilities directly. As stated previously, it is felt that the expert will respond more meaningfully if he is exposed to a systematic interviewing process than if he is asked to make direct assessments.

2. The appraiser is a knowledgeable expert in the area of concern.

3. The expert has been instructed in the basic concepts of probability theory. He must understand the meaning of "probability of occurrence" as employed in this section, as well as a conception of probability density functions and distribution functions. We contend that there is a direct relationship between the expert's knowledge of probability theory and his facility in responding to the analyst's questions.

2.5 THE STANDARD LOTTERY

2.5.1 Introduction.

The objective of the Standard Lottery technique is the derivation of a probability density function over all possible values of a given component characteristic. Again, the procedure involves presenting the expert with two gambling situations. This technique differs from the two previously discussed methods in that it does not involve the process of varying actual probabilities or performance levels until indifference is achieved. Instead, numbers representing randomly selected lottery tickets from a batch of 100 are varied in an attempt to achieve indifference. In essence, the number of such tickets directly infers component risk probabilities.

2.5.2 Description.

The technique is based on the following basic lottery description. In a lottery a contestant pruchases as many tickets as desired. The more tickets he purchases, the greater his chance of winning the contest prize. After the purchase of tickets is completed, one number is randomly drawn from a lot of 100 equally likely numbers. That is, each contestant fully understands that any number between 1 and 100 has an equal chance of being selected. The winning contestant is that individual who owns a lottery ticket with the number on it coinciding with the number selected. For example, a contestant might have randomly purchased 40 individual tickets out of the lot of 100 tickets. If one of his tickets coincided with the number chosen, he would be the winner. Before the drawing, this contestant knows that he has 40 chances out of 100 of holding the winning ticket regardless of his method of selecting the purchased tickets. That is, he does not feel that it would be worth the slightest effort to select tickets with particular numbers on them; the probability that a lottery ticket in the group numbered 1 to 40 will be the winning ticket, is equal to the probability that a ticket in a randomly selected group of 40 tickets will be the winning ticket.

The standard lottery technique of eliciting a subjective probability density function employs the foregoing concept of a lottery. In this method, the expert is presented with a hypothetical lottery of 100 lottery tickets (numbered from 1 to 100). This lottery is to be used as a standard of comparison in helping the expert decide what probability value to assign to the possible realization of a given component characteristic level. The questioning procedure is as follows:

a. Specify a possible component characteristic value (e.g., thrust = 36,000 lb) for the real-world event.

b. Direct the expert to imagine that he is given a choice between a certain number of tickets in the standard lottery with a prize of value V (e.g., $10) and the right to receive the same prize if the real-world event (i.e., thrust = 36,000 lb) is realized.

c. For a given initial number of lottery tickets (e.g., 30), ask the expert which alternative gamble he feels has the greatest chance of winning the prize: (1) the realization of the real-world event, or (2) the holding of the specified number of tickets (i.e., 30) of a lottery of 100 tickets outstanding.

d. If one gamble is preferred over the other, next vary the given number of tickets (i.e., increase if the expert chooses the real-world event in step c, decrease if he chooses the lottery alternative) and repeat the questioning of step c.

e. Repeat steps c and d until, in his opinion, the expert feels the possibility of receiving the prize (e.g., $10) in the event of, for example, thrust = 36,000 lb, has exactly the same likelihood as, say, 70 tickets in the standard lottery. Thus, it is inferred that the expert considers these two events equally likely. Thus, he assigns a probability weight of 0.7 (70/100) to the event thrust = 36,000 lb.

f. Employing steps a through e, the expert can proceed analogously to assign probability weights to all other possible real-world events.

Resulting from the questioning procedure above, a hypothetical example of the final probability density table is shown below.

TABLE 2.12 FINAL PROBABILITY DENSITY

Thrust (lb)	No. of Tickets	Implied Probability
32,000	0	0.0
34,000	10	0.1
36,000	70	0.7
38,000	20	0.2
40,000	0	0.0

The weights in the table above must, of course, ultimately sum to 1.0. To achieve this, it will sometimes become necessary to repeat the decision process and/or normalize the resulting probability values. Hence, the table represents the placement of collectively exhaustive and mutually exclusive events (e.g., thrust values) into a one-to-one correspondence with a set of collectively exhaustive and mutually exclusive events in the standard lottery.

2.5.3 Advantages, Limitations, Assumptions.

Again, this technique provides an improved process for eliciting subjective responses over direct interrogation. It is similar to the Choice-Between-Gambles Technique for deriving probability density functions and thus offers the same advantages, limitations, and assumptions with one addition.

Since probability statements are not made directly, the expert with little probability theory background may be more comfortable with this technique.

Of course, the success of the technique also depends upon the engineers familiarity with the lottery-type betting situation.

2.6 THE MODIFIED CHURCHMAN-ACKOFF TECHNIQUE

2.6.1 Introduction.

The Modified Churchman-Ackoff Technique differs from those described heretofore in the following ways: (1) it does not involve betting situations, (2) the expert is not asked to reveal indifference values of the parameter in question, and (3) the expert is instead asked to make "greater than," "equal to," or "less than" evaluations regarding relative probabilities between two sets of values and relative probability assessments with respect to the most probable characteristic value. The resulting relative probability scale is easily transformed into a probability density function and subsequently into a distribution function.

2.6.2 Description.

In this technique, the expert must reveal a range of possible values which the component characteristic could possibly realize. Employing perhaps the Choice-Between-Gambles method of deriving a density function, end point values of zero probability of occurrence must be specified. These values need only be any low and high values which the expert specifies as having zero probability of occurrence in the proposed system.

Next, individual values within the range of possible values must be determined. These values, which will form the set of comparative values for this technique, are specified by the following approach:

(1) Start with the smallest.

(2) Progress upward on the scale of values until the expert is able to state a simple preference regarding the relative probabilities of occurrence of the two characteristic values. If he is able to say that he believes one value has either a greater chance or a lesser chance of occurring than the other of the two values, then it is inferred that the expert is able to discriminate between the two values.

(3) Using the higher of the two previously specified scale values as a new basis, repeat step (2) to determine the next value on the scale.

(4) Repeat steps (2) and (3) until the high end point value of the range of parameter values is approached.

Employing this procedure for the aircraft example, one might obtain the results in Table 2.13

TABLE 2.13 CHARACTERISTIC VALUES FOR THRUST EXAMPLE

Θ_1 = 35,000 lb

Θ_2 = 36,000

Θ_3 = 37,500

Θ_4 = 38,500

Θ_5 = 40,000

Θ_6 = 41,000

Θ_7 = 41,500

The descending order of probability of occurrence can be determined by applying the following paired comparison method.

Ask the expert to compare, one at a time, the first discrete value (Θ_1) of the set to each of the other values (Θ_2, Θ_3, etc.), stating a preference for that value in each group of two values that he believes has the greater chance of occurring (denoting a greater probability of occurrence by >, and equal chance by =, and a lesser chance by <). The following hypothetical preference relationships could result for a set of 7 values ($\Theta_1 < \Theta_2$, $\Theta_1 < \Theta_3$, $\Theta_1 < \Theta_4$, $\Theta_1 < \Theta_5$, $\Theta_1 < \Theta_6$, $\Theta_1 = \Theta_7$).

Next, ask the expert to compare, one at a time, the second discrete value (Θ_2) of the set to each of the other values succeeding it in the set (i.e., Θ_3, Θ_4, etc.). The following preference relationships might result ($\Theta_2 < \Theta_3$, $\Theta_2 < \Theta_4$, $\Theta_2 < \Theta_5$, $\Theta_2 > \Theta_6$, $\Theta_2 > \Theta_7$).

Continue the process until all values (Θ_i) have been compared to the others. For example Table 2.14 lists preferences which might result for the remaining thrust values.

TABLE 2.14 PAIRED COMPARISONS

Θ_3 vs $\Theta_4,\ldots,\Theta_7$	Θ_4 vs $\Theta_5,\ldots,\Theta_7$	Θ_5 vs Θ_6, Θ_7	Θ_6 vs Θ_7
$\Theta_3 < \Theta_4$	$\Theta_4 > \Theta_5$	$\Theta_5 > \Theta_6$	$\Theta_6 > \Theta_7$
$\Theta_3 > \Theta_5$	$\Theta_4 > \Theta_6$	$\Theta_5 > \Theta_7$	
$\Theta_3 > \Theta_6$	$\Theta_4 > \Theta_7$		
$\Theta_3 > \Theta_7$			

Now total the number of times (Θ_i) value was preferred over other values. The results for this procedure are listed in Table 2.15.

List the values in descending order of simple ordinal probability preference and change the symbols for each value from Θ_i to X_j as shown in Table 2.16.

TABLE 2.15 SUMMARY OF PREFERENCE RELATIONSHIPS

Θ_4 = 6 times

Θ_3 = 5 times

Θ_5 = 4 times

Θ_2 = 3 times

Θ_6 = 2 times

Θ_1 = 0 times

Θ_7 = 0 times

TABLE 2.16 TRANSFORMATION

Characteristic Value (lb)		Preference Rank	New Symbol
38,500	Θ_4	1	X_1
37,500	Θ_3	2	X_2
40,000	Θ_5	3	X_3
36,000	Θ_2	4	X_4
41,000	Θ_6	5	X_5
35,000	Θ_1	6	X_6
41,500	Θ_7	7	X_7

Arbitrarily assign a rating of 100 points to the characteristic value with the highest subjective probability (e.g., X_1). Then, as in the first step, question the expert regarding the relative chance of occurrence of each of the other values on the ordinal scale in Table 2.16 with respect to the value at the top of the scale. Assigning X_1 a rating of 100 points, the expert is first interrogated as to his feeling of the relative chance of occurrence of the second highest scale value (e.g., X_2), with respect to X_1. Does it have 25 percent chance? 60 percent? 70 percent? 80 percent? As much chance of realization as X_1? The relative probability rating, based on 100 points, (i.e., 100 percent as much chance) will then be posted for X_2.

Next, question the expert about the relative chance of occurrence of the next highest scale (e.g., X_3) first with respect to the most preferred value (X_1), and second with respect to the second most preferred scale value (X_2). The resulting numerical ratings should concur. For example, if the expert decides that X_2 has 8/10 as much chance of occurring as does X_1, the ratings become X_1 = 100 points and X_2 = 80 points.

If the expert expresses a belief that X_3 has 1/2 as much chance as X_1 and 5/8 as much chance as X_2 (as a validity check), this confirms that the relative probability of occurrence rating for X_3 is 50, and the scale becomes X_1 = 100 points, X_2 = 80 points, and X_3 = 50 points.

Continue the process for each remaining successively lower scale value on the ordinal scale shown in Table 2.16. Determine the relative number of points to be accorded each value with respect to the top scale value and with respect to all other values on down the scale which are above the characteristic value in question.

In the event of minor disparities between relative probability ratings for a given value, the average of all such ratings for that characteristic value might be computed. For example, X_4 might be determined to be 3/10 as probable as X_1, 1/4 as probable as X_2, and 1/2 as probable as X_3. The 3 absolute ratings for X_4 are thus inferred to be 30, 20, and 25 points respectively. The average of these ratings is 25. However, before averaging such figures, it might be beneficial to have the expert reevaluate his relative ratings for X_4 with respect to X_1, X_2, and X_5.

As a result of the above process, the relative probability values shown in Table 2.17 might be attained.

TABLE 2.17 RELATIVE PROBABILITY RATINGS

RX_1 = 100 probability points

RX_2 = 80 probability points

TABLE 2.17 (Continued)

RX_3 = 50 probability points

RX_4 = 25 probability points

RX_5 = 10 probability points

RX_6 = 0 probability points

RX_7 = 0 probability points

Finally, the scale of relative probability values can be converted directly into a scale of actual probability density values by letting $P(X_1)$ equal the actual subjective probability of occurrence of the highest value. Then, $P(X_2)$ is then defined as

$$\frac{R(X_2)}{R(X_1)} \; [P(X_1)] \quad .$$

Similarly $P(X_i)$ is defined as

$$\frac{R(X_i)}{R(X_1)} \; [P(X_1)]$$

for i = 2, 3,...,7.

Assuming that the independent characteristic values evaluated represent all possible values attainable by the component characteristic, the respective probabilities must sum to 1.0 (i.e., $P(X_1) + P(X_2) + P(X_3) + P(X_4) + P(X_5) + P(X_6) + P(X_7) = 1.0$). Substituting the expressions for $P(X_i)$, i = 2,...,7, it follows that

$$P(X_1) + \frac{R(X_2)}{R(X_1)} [P(X_1)] + \frac{R(X_3)}{R(X_1)} [P(X_1)] + \frac{R(X_4)}{R(X_1)} [P(X_1)]$$

$$+ \frac{R(X_5)}{R(X_1)} [P(X_1)] + \frac{R(X_6)}{R(X_1)} [P(X_1)] + \frac{R(X_7)}{R(X_1)} [P(X_1)] = 1.$$

Solving this equation for $P(X_1)$, the remaining $P(X_i)$, i = 2,...,7 can be determined using the relationship

$$P(X_i) = \frac{R(X_i)}{R(X_1)} [P(X_1)] \quad .$$

As an illustration, consider the relative probability ratings in Table 2.17. Using these values, the preceding equation is given by

$$P(X_1) + \frac{80}{100} P(X_1) + \frac{50}{100} P(X_1) + \frac{25}{100} P(X_1) + \frac{10}{100} P(X_1) = 1 \quad .$$

Solving this equation, $P(X_1) = 0.377$.

This value can be used to determine the remaining probabilities as follows:

$$P(X_2) = \frac{RX_2}{RX_1} P(X_1) = 0.80(0.377) = 0.301$$

$$P(X_3) = \frac{RX_3}{RX_1} P(X_1) = 0.50(.0377) = 0.189$$

$$P(X_4) = \frac{RX_4}{RX_1} P(X_1) = 0.25(0.377) = 0.095$$

$$P(X_5) = \frac{RX_5}{RX_1} P(X_1) = 0.10(0.377) = 0.038$$

$$P(X_6) = \frac{RX_6}{RX_1} P(X_1) = 0(0.377) = 0.000$$

$$P(X_7) = \frac{RX_7}{RX_1} P(X_1) = 0(0.377) = 0.000$$

The resulting probability density appears in Table 2.18.

TABLE 2.18 PROBABILITY DENSITY

Component Characteristic Value	Probability
X_1	0.377
X_2	0.301

TABLE 2.18 (Continued)

Component Characteristic Value		Probability
X_3		0.189
X_4		0.095
X_5		0.038
X_6		0.000
X_7		0.000
	TOTAL	1.000

2.6.3 Advantages.

This technique offers an alternative to previous methods of eliciting absolute subjective probability responses. In this case relative probabilities with respect to one chosen most probable characteristic value are derived. In some situations, the expert may think it easier to make evaluations with respect to a characteristic state that he feels has the greatest possibility of realization. However, at this writing the technique remains to be empirically tested in this capacity.

2.6.4 Limitations.

In addition, this technique offers a systematic method of checking the consistency of relative value judgments made by the experts. This enhances the validity of the resulting probability distribution.

The technique does not involve betting situations which are generally considered more successful in eliciting correct responses. Instead, it involves an untested approach of directly eliciting relative percentage chances of occurrence statements for each value with respect to the occurrence of other characteristic values (e.g., does a thrust of 35,000 lb have half as much, or 70 percent, or 90 percent as much chance of occurring as 38,500 lb?).

As with the other techniques of this section, the probability values are still judgments. This is, of course, the limitation of all techniques involving subjective (as opposed to objective) decision making.

2.7 THE DELPHI PROCEDURE

2.7.1 Introduction.

The preceding techniques delineated in this section have been directed at decision-making situations in which individual experts are interrogated regarding subjective probability preferences. The resulting functions were therefore assumed to be based solely on that expert's knowledge, experience, and intuition. They are also assumed to be void of external influence from other individuals. However, in many situations there may exist a group of experts and a group probability density function may be sought instead of individual density functions. This is based on the old saying "two heads are better than one" or more generally "n heads are better than one." But how does one draw upon this group judgment to estimate the group probability density function?

Historically, the approach for obtaining a group consensus has been the formation of committees, commissions, or councils. While the basic philosophy may be sound, committees tend to pressure individuals into conforming. The pressure to conform may not be applied directly, but participants are certainly aware of this pressure. In addition, all opinions may not be expressed because of the personalities of the individuals and/or because of the relationship of the individuals within the group.

Another drawback of the committee is its tendency to spend a great deal of time discussing irrelevant issues. While the potential for inefficiency is obvious, there is also the possibility that the irrelevant information may degrade the groups opinion.

More serious than any of the preceding potential problems is the possibility of a complete breakdown of the committee. Breakdown here refers to an inability to arrive at a general consensus of opinion. Consider the situation where there are several conflicting opinions. The supporters of these opinions may get emotionally involved in defending their positions and loose sight of the objective of the committee. Even when an agreement is reached, it will probably not represent the consensus of the group.

The Delphi procedure is an alternative to the committee approach for eliciting a group judgment. It "attempts to improve the panel or committee approach in arriving at a forecast or estimate by subjecting the views of individual experts to each other's criticism in ways that avoid face-to-face confrontation and provide anonymity of opinions and of arguments advanced in defense of these opinions. In one version, direct debate is replaced by the interchange of information and opinion through a carefully designed sequence of questionnaires. The participants are asked not only to give their opinions but the reason for these opinions, and, at each successive interrogation, they are given new and refined information, in the form of opinion feedback, which is derived by a computed consensus from the earlier parts of the program. The process continues until further progress toward a consensus appears to be negligible. The conflicting views are then documented. (Reference 3).

The primary features of Delphi procedures are:

a. Anonymity of the source of information among experts.

b. Iteration with controlled feedback of group responses from iteration to iteration.

c. Statistical group response (the prescribed measure is the median) (Reference 4).

It should be emphasized that the Delphi procedure has broader potential in the analysis of uncertainty than in estimating a group probability density function. However, since this is a subjective probability section, this application will be discussed first.

Before continuing with the description of the application of the procedure in estimating a group probability density function, some experimental results will be presented. These experiments provide indications that the procedure has merit and deserves consideration.

2.7.2 Experimenting.

The results of three RAND experiments will be discussed briefly. The objectives of these experiments were:

1. Examine "whether the use of iteration and controlled feedback have an advantage over the mere statistical aggregation of opinions." (Reference 4).

2. Compare "the performance of groups using face-to-face discussion with groups employing anonymous, questionnaire-feedback interaction." (Reference 4).

3. Examine whether the selecting of subgroups based on self-ratings will produce better results when incorporated in the Delphi procedure (Reference 5).

---

[3] Quade, E. S., and Boucher, W. I.; Systems Analysis and Policy Planning, Applications in Defense, Elsevier Publishing Co., New York, 1968, p 334.

[4] Dalkey, N.; An Experimental Study of Group Opinion, The Delphi Method, Vol 1, No. 5, September 1969, Futures, p 408.

[5] Dalkey, N., Brown, R., and Cochran, S.; Use of Self-Ratings to Improve Group Estimates, Experimental Evaluation of Delphi Procedures, Technological Forecasting 1, 1970, p 283.

In all of these experiments, almanac type questions were used. These type questions were used because the participants probably didn't know the answers, but they could make an informed guess. In addition, since the answers were known, improvement in responses could be measured and responses could be compared.

1. The responsive results of these experiments were:

a. "On the initial round a wide spread of answers typically ensued.

b. "With iteration and feedback, the distribution of individual responses progressively narrowed (Convergence).

c. "More often than not, the group response (defined as the median of the final individual responses) became more accurate." (Reference 4).

2. There were two parts to the second experiment mentioned previously. In the first part, "the basic result was that the median response of the questionnaire group was more accurate in 13 cases and the consensus of the face-to-face group was more accurate in 7 cases. Considered as an isolated experiment, this result is not statistically significant. However, when this experiment is considered with several others showing the same kind of outcome, the results appear more significant." (Reference 4). In the second part of the experiment, face-to-face interaction was shown to have an overall degrading effect on the groups responses. These results are also in agreement with other experiments in which Delphi participants were put into groups after participating in a Delphi exercise. In all instances, the responses in the face-to-face group were degraded over the Delphi responses. (Reference 4).

3. More accurate subgroups can be selected for many questions. "In addition, answers to the remaining questions improve upon feedback, so that a combination of subgroups selection and feedback produces a significantly larger number of improved group responses than could be obtained by feedback alone." (Reference 5).

For a detailed description of these experiments and their results the interested reader should consult the above references.

What do these results mean? Certainly, they provide some reasonable indications that anonymous iteractions with controlled feedback has merit. In addition, there are indications that Delphi may

---

[4] Ibid., pp 415-418.

[5] Ibid., p 283

provide a better method for obtaining a group judgment and the use of self-rating with the Delphi procedure may produce better results. However, experiments such as these can never provide more than an indicator because they lack some of the critical characteristics of real-world situations, such as:

a. If the answers are not reasonable, the consequences are not important.

b. The participants have no responsibility.

c. Real-world uncertain situations can probably never be measured for accuracy.

d. Real-world questions are generally value oriented.

e. The respondents are not experienced in forming opinions (i.e., the respondents are not representative of the type of individuals who would be real-world participants) and aren't really experts.

f. The subject matter is not realistic (i.e., too many diverse areas).

Of course, it is recognized that realistic experiments are probably not possible. However, these comments are directed at those individuals who might be inclined to draw strong inferences about real world applications from these types of experiments.

In summary, these experimental results provide some good indicators and valuable insight into the Delphi procedure, but they do not provide conclusive evidence.

The next two sections will be devoted to a description of two different applications of the Delphi procedure.

2.7.3 Application of the Delphi Procedure for Estimating a Group Probability Density Function.

The steps of the procedure for estimating a group probability density function are outlined below for the jet engine thrust example.

Employing the first two steps of the modified Churchman-Ackoff technique, each expert is asked to reveal his estimate of the total range of values which the component characteristic could realize and the individual values within this range which form the sets of comparative values. Then, list all values specified by all of the experts. These will form the list of values to be investigated. For example, Table 2.19 lists the responses of five experts concerning the values of thrust which they judged could be discriminated between on a probability basis.

TABLE 2.19 POSSIBLE CHARACTERISTIC VALUES BY EXPERT

Expert 1 (lb)	Expert 2 (lb)	Expert 3 (lb)	Expert 4 (lb)	Expert 5 (lb)
28,000	29,000	30,000	28,000	30,000
31,000	31,000	32,000	30,000	32,000
32,000	33,000	34,000	32,000	34,000
34,000	36,000	37,000	33,000	36,000
37,000	40,000	39,000	36,000	37,000

The list of characteristic values to be investigated are included in Table 2.20.

TABLE 2.20 LIST OF POSSIBLE DISCRETE CHARACTERISTIC VALUES

Y_1	28,000 lb
Y_2	29,000
Y_3	30,000
Y_4	31,000
Y_5	32,000
Y_6	33,000
Y_7	34,000
Y_8	36,000
Y_9	37,000
Y_{10}	39,000
Y_{11}	40,000

In the first round, randomly select a characteristic value from the list in Table 2.20 and ask each expert to give an independent estimate of its probability. One of the previously described techniques might be employed here as a tool in eliciting each expert's subjective probability response. Each expert is questioned alone. In addition, interrogate each expert regarding his reasons for evaluating the characteristic value as he did.

Arrange the probability responses from all experts in order of magnitude, and determine its quartiles, Q_1, M, Q_3, so that approximately one quarter of all estimates lie in each interval. For example, for the selected thrust of 31,000 lb, the probabilities for experts E1, E2, E3, E4, and E5 might occur as shown in Figure 2.2.

Reveal the values and responses of each interval to each member, and if his estimate lies outside the first round interquartile range, Q_1, to Q_3, ask him to state his reasons why the answer should be lower (or higher) than that of the 75 percent majority opinion expressed in the first round.

Give these new responses back to all respondents by communicating the new range of the new quartile values, along with independently stated reasons for the estimates laying outside the 75 percent majority opinion. The experts are now asked to consider the reasons given, weigh their feasibility, and revise their own previous estimates accordingly. For the thrust example, assume the second round scale is as shown in Figure 2.3.

If newly revised probabilities still fall outside the second round interquartile range, respondents are asked to state why they found previous arguments unconvincing enough to draw them toward the median.

In the third round, the quartile results of round 2 are submitted to respondents along with the counter arguments elicited. These respondents are then asked to make a final revision of their estimates.

The mean value of the resulting round 3 estimates is taken as the group response as to what the subjective probability consensus for the thrust value should be. For the thrust example, the mean third round subjective probability estimate for thrust = 31,000 lb is

$$[2(0.275) + 2(0.300) + (0.325)]/5 = 0.295 = P(\text{Thrust} = 31{,}000 \text{ lb}) \; .$$

Now, repeat the procedure for a second possible characteristic value. Normalize the distribution if necessary.

2.7.4 Other Applications.

Up to now, the analysis of uncertainty has been restricted to the context where the analysis was part of a much broader analysis. Further, it was assumed that the problem was well defined and the

EXPERT	E_4	E_2	E_1	E_5	E_3
PROBABILITY	.15	.25	.3	.325	.4

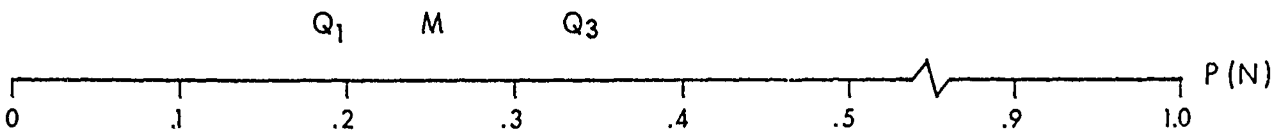


Figure 2.2 Probability Responses: 1ST Round.

EXPERT	E_4	E_2	E_1	E_5	E_3
PROBABILITY	.25	.275	.3	.325	.35

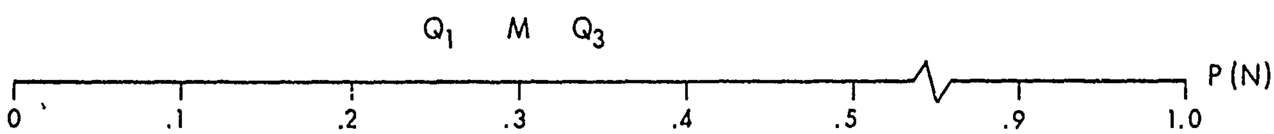


Figure 2.3 Probability Responses: 2ND Round.

EXPERT	E_2	E_4	E_1	E_5	E_3
PROBABILITY	.275	.275	.3	.325	.325

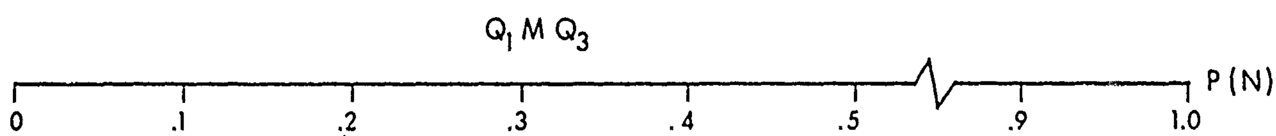


Figure 2.4 Probability Response: 3RD Round.

uncertain variables or areas had been adequately identified. However, risk analysis includes more than consolidation of group information in order to evaluate the risks, although the Delphi procedure can be used in this manner. Consider the following questions:

a. What are the alternatives?

b. What is the best alternative?

c. What should be the future goals of an organization?

d. What is the most desirable policy?

e. What is the magnitude of the non-nuclear threat in Europe in 1980?

f. What are the variables or areas of interest?

The answers to these type questions definitely are uncertain, but the interest here is not in the explicit treatment or uncertainty. More specifically, one is interested in considering the uncertainty implicit in answering these type questions. These types of questions address problem definition and situations where quantitative models are inadequate. The models may be inadequate because the output measure is not well defined (e.g., how does one measure desirability?), or the measure may be a function of many variables, and the relationship between these variables may be unknown. In any event, these situations do exist and must be analyzed.

In addition, the last question demonstrates that Delphi procedures might be used in the identification procedure in risk analysis. However, the reader is cautioned that this is not always the case, and this procedure is no substitute for systematically investigating the program. Without a good background, one would not be able to structure the exercise or analyze the results meaningfully.

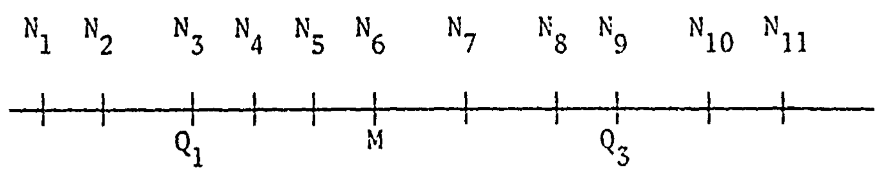
Usually there will be a group of knowledgeable people who can be drawn upon to obtain answers to these types of questions. Since there is generally no way of selecting one expert's opinion over another, a group consensus of opinion is generally sought. Once again, the analyst is confronted with the choice between the Delphi procedure and the committee approach. Which is the best way? No one can answer that question for sure. It depends on the particular situation. This selection problem will be discussed in more detail later.

To illustrate that the Delphi procedure can be used in these types of situations, two examples are presented in the next section.

2.7.5 Examples.

In this section, two examples of the Delphi procedure are presented. In the first example, the procedure is used to estimate a quantity such as the world population in 1985. In the second example, a much broader policy question is addressed. It is hoped that these two examples will serve to clarify the mechanics and underlying principles of the technique. Both examples have been extracted verbatim from Systems Analysis and Policy Planning, Applications in Defense by Quade and Boucher.

> "Example: Choosing a Number by Delphi. Consider the common situation of having to arrive at an answer to the question of how large a particular number N should be. (For example, N might be the estimated cost of a measure, or a value representing its over-all benefit.) We would then proceed as follows: First, we would ask each expert independently to give an estimate of N, and then arrange the responses in order of magnitude, and determine the quartiles, Q_1, M, Q_3, so that the four intervals formed on the N-line by these three points each contained one quarter of the estimates. If we had eleven participants, the N-line might look like this:



> Second, we would communicate the values of Q_1, M, Q_3 to each respondent, ask him to reconsider his previous estimate, and, if his estimate (old or revised) lies outside the interquartile range (Q_1, Q_3), to state briefly the reason why, in his opinion, the answer should be lower (or higher) than the 75-per cent majority opinion expressed in the first round. Third, we would communicate the results of this second round (which as a rule will be less dispersed than the first) to the respondents in summary form, including the new quartiles and median. In addition, we would document the reasons that the experts gave in Round 2 for raising or lowering the values. (As collated and edited, these reasons would, of course, preserve the anonymity of the respondents.) We would then ask the experts to consider the new estimates and the arguments offered for them, giving them the weight they think they deserve, and, in light of this new information, to revise

their previous estimates. Again, if the revised estimate fell outside the second round's interquartile range, we would ask the respondent to state briefly why he found unconvincing the argument that might have drawn his estimate toward the median. Finally, in a fourth round, we would submit both the quartiles of the third distribution of responses and the counterarguments elicited in Round 3 to the respondents, and encourage them to make one last revision of their estimates. The median of these Round 4 responses could then be taken as representing the group position as to what N should be.

"Example: Policy Advice from Delphi. The Delphi technique can also be applied to broad policy problems. For example, let us consider how it might be used to uncover and evaluate measures that might help to speed recovery of a nation after a thermonuclear war.

"There are a number of reasons why an approach to this problem via the development of a mathematical model or a computer simulation might not be the most desirable way to proceed. If we had in mind six or eight fairly well-defined and promising alternative postwar measures, we might consider adding a "recovery" model to one of the many models that have been constructed to compute the damage caused by a nuclear attack. Assuming this could be done, the alternatives could then be compared in the traditional way used for comparing alternative force structures, employing a range of different war initiation scenarios and undertaking sensitivity analyses of the uncertain parameters.

"But the concept of 'recovery' is not very well defined. Very few of the many measures that might aid the survival of a nation or an area after a thermonuclear attack have been studied extensively. The emphasis so far has fallen primarily on measures such as shelters and active defense, which seek to reduce the immediate effects of the attack, rather than on measures to speed recovery after the initial effects of an attack have been experienced. Almost everyone has ideas about recovery measures of this type that might be helpful, but seldom any well-developed notion of their relative effectiveness and cost. Thus there is a need to survey these ideas - to create an atmosphere in which they may be brought forth, subjected to critical review, modified and ordered according to various criteria with respect to their possible effectiveness, acceptability, and costs, including social costs. The Delphi technique is well-suited to this task.

"In addition to the presence of so many ill-defined alternatives, and the difficulties with the notion of recovery, there are a number of other reasons why an approach to the problem that puts emphasis on informed judgment is desirable. The decisionmakers who would use the study would clearly be in the best position to judge the acceptability of measures that might either require radical changes in the prewar way of life or imply such changes for the postwar period - for example, how far to violate the rights of privacy or favor one sector of the economy or country over another if nuclear war were to come. But their decisions would necessarily be based on many lowly but important relationships that require the intuition and judgment of specialists. Determining objectives what we want to accomplish in the way of recovery and how we might distinguish one type of postwar world from another - must also be the responsibility of the decisionmaker. But how to attain these objectives would require contributions from many disciplines.

"The alternative provided by the Delphi technique is to give up for the moment any attempt to compute the state of the postwar environment at various times after hostilities have ceased and instead to try simply to rank alternative prewar policies on the basis of the qualities that promise, in the judgment of specialists, to contribute the most to postwar recovery. This procedure cannot demonstrate beyond all reasonable doubt that a particular course of action is best. At most, it can assess some of the implications of choosing certain alternatives over others. But the systematic searching out and partial ordering of promising steps could be extremely valuable.

"We should be under no illusion that for this problem a Delphi procedure would be the easiest thing in the world to carry out. In order to persuade the proper people to authorize or to participate in such a study, the following points would have to be brought to their attention. One, the effort would not be intended as a substitute for other research. Two, if nothing else, it would highlight areas needing detailed study and in general, stimulate further work. Three, ideas provided in the course of the study - because of their possible half-baked character - would be kept anonymous unless attribution was specifically authorized. And four, the entire effort, in terms of manpower, could be kept quite minor, even though as much as ten months might be needed to complete the study, since getting responses to questionnaires is just slow business.*

"Since the kind of survey being proposed is not a statistical survey of the Gallup type, but an attempt to generate ideas and to use the respondents to trace out the interrelationships among these ideas and the consequences of their adoption, it is immaterial whether the respondents form a representative sample of the initially known points of view. What matters is that the viewpoints of persons with all major relevant backgrounds have a chance of being voiced.

"Assuming that our study would involve a range of experts both within and outside the organization conducting it, the respondents might be organized into several "units," so that the administrative task of running the experiment could be kept simple. Each unit might consist of a central committee of three plus a panel of six to twelve respondents. The committee chairman would be the person responsible for organizing his unit's activity, for maintaining liaison with the project director, and for transmitting the responses of his unit. One or more units might be located within the organization carrying out the study and the other units at some of the various places where there is a concentration of respondents. Alternatively, the respondents might be dealt with directly or split into functional groups or disciplines such as ecology, economic growth, and so on.

"The inquiry itself could be broken down into four to six successive rounds, each based on a suitably formulated questionnaire. Only round one would necessarily involve all respondents.

---

*Incidentally, there exists an Act of Congress (5 U.S.C. Sec. 139, c-e [1942]) that forbids a government agency to conduct or to sponsor a study in which identically worded questionnaires are circulated to more than nine respondents without prior permission of the Bureau of the Budget. Since the intent of the Act is to keep businessmen from being bothered with a continuous stream of government forms - not to hamper scientific investigation - users of the Delphi technique whose support comes from government funds should not have difficulty obtaining such permission. Of course, one could confine the respondents (except for at most nine outsiders) to the research organization (this includes consultants) or the sponsoring agency.

"The first questionnaire would contain, in addition to the questions themselves, a brief background statement explaining the purpose of the study. It would include a statement that responses will be handled anonymously, except that approval for the use of names may eventually be asked in case certain suggestions are deemed worthy of being recommended for further action. Only the members of the steering committee would initially be cognizant of the authorship of ideas. In the statement suggestions would be included about keeping the proposals in practical operational terms and avoiding generalities. The respondents would be urged to include all suggestions that they think should be examined, even though they might be dubious about advocating them.

"The following sample questionnaire incorporates a number of these suggestions. Since it is addressed more to the readers of this book than to potential respondents, considerable reworking would be required before it could actually be used.

Questionnaire 1

This questionnaire is being submitted to you in an effort to elicit fresh ideas on what steps should be taken to reduce the problem of postattack recovery after a thermonuclear exchange. We are not looking for measures that reduce the number of weapons impacting (ABM, for example) or measures that reduce their efficiency (such as shelters). Primarily we are looking for ways to help restore agriculture and manufacturing and the structure of society and government. An earlier study has suggested that the measures we are seeking to identify and weigh fall into three classes: preventive, which would aim at reducing the damage to our resources, such as food stocks and water and power sources; emergency, which would attempt to deal with the distribution and management of supplies to sustain the population after the war; and long run, which would deal with recovery proper. Regardless of your feelings about the probability of nuclear war and the futility of such actions - in themselves or in contrast to the results we might obtain if we contributed equal resources to deterrence - ask yourself what measures should be considered.

This effort is being conducted very much in the spirit of a brainstorming session, except that it sets out to collect ideas in written form rather than through the give-and-take of open debate. At this stage, therefore, it would be entirely in order for you to submit ideas even if you yourself consider them half-baked, or if you merely regard them as worthy of further exploration without wishing to endorse them, or if they would only gain full meaning within an adequately elaborated context. Remember that this survey is in no way intended as a substitute for other research; indeed, its chief virtue might be to highlight areas needing detailed study and, in general, to stimulate further work.

Question A. If you were a close advisor to the President, what actions would you advise him to consider taking (including recommentation of legislation to Congress) that might speed recovery after a thermonuclear attack?

The following considerations - the list is by no means complete - seem relevant to this question. You may wish to delete or modify some items or add others. They are offered only to spark thought, and are listed randomly to avoid prejudging the order of importance or the feasibility of any measures.

1. Since the control of infectious diseases could be a serious problem in the disrupted postattack environment, should current public health policies be reviewed for possible changes that would improve their effectiveness in a postattack situation? What policies? What changes?

2. A number of studies indicate that fires, both urban and wildland, as well as their sequelae of floods, erosion, and additional fire hazards, could be serious long-term problems in the postattack environment. Is there a need to review current fire prevention and control practices for possible changes and innovations that could improve our postattack capabilities to cope with these problems? What changes might be made? It has been suggested, for example, that we might undertake controlled burning prewar and also create appropriate firebreaks to prevent wildland fires from encroaching on contiguous urban areas or to keep urban fires from spreading to the countryside. We might also consider some steps to provide for re-seeding burned areas postattack to reduce erosion and flooding.

3. How serious a problem would it be to find feasible alternative postattack land uses that would be keyed to postattack requirements for food and other agricultural products? For example, what other crops could be grown on land too heavily contaminated with fallout to grow food, or what food crops could be grown on land not heavily contaminated but now used to grow non-food products?

4. What priorities should be observed in restoration of facilities postattack?

5. Should differential protection be provided for different segments of the population?

6. Is organizational damage likely to be a serious problem in the postattack environment?

Question B. What research should be undertaken by the scientific and technical community that might either lead to or accelerate the discovery of measures that would help speed postwar recovery?

Again, here are a few possibilities that you may wish to consider in your response.

1. Develop models. It might, for example, be important to build a flexible modular fallout model, or a model of the ignition and spread of urban fire and its impact on population in the fire area, including the protection afforded by available shelters against heat and carbon monoxide poisoning. A model of wildland fire that would relate ignition and spread to plant cover, season of year, weather, geographical region, and the nature of the nuclear attack might also be useful, as would models of a disrupted economy, since current models all seem to assume an organized society.

2. Perform further research. Research in atmospheric physics, for example, might give us a way to estimate the the effects of nuclear exchanges on weather and climate. Similarly, research might be undertaken on ecological disturbance or on the long-term genetic effects of radiation on man. (Both of these problems have already been studied in some detail, but much ignorance remains.)

3. Develop technologies for food storage and synthesis.

4. Develop contingency plans for priorities in resource allocation by age, by sector of the economy, or by some other standard.

"Once the responses to this first questionnaire had been received, the next, and hardest, step would be for the steering committee to sort and collate them, clarifying their meaning through checks with the respondents if necessary, eliminating obviously non-operational suggestions, doing some minor editing and, hopefully, generating useful additions to the list.

"The list of proposals thus produced might then be submitted either directly to the original respondents or, as an intermediate step to obtain further refinement, to the "unit" committees. The result of this review might be the elimination of, say, two-thirds of the proposals as being less promising. The remainder would then be annotated by the steering committee with brief arguments pro and con; they might also be ranked by merit according to some consensus formula.

"Because the wording of every questionnaire but the first depends on the outcome of preceding rounds, we can at best indicate only the form the remaining questionnaires might take. The second might look something like this:

Questionnaire 2

The tabulation given below contains a list of tentative proposals to speed postwar recovery. We would like you to give us your judgment of each item in terms of its desirability, its feasibility, and its potential importance (assuming feasibility).

For each item, check one box under Columns A, B, and C. In making this evaluation, consider the intrinsic rather than relative merits of the proposal.

		A Desirability					B Feasibility					C Importance			
No.	Proposal	Desirable	Mildly Desirable	Doubtful	Mildly Undesirable	Undesirable	Definitely Feasible	Possibly Feasible	Doubtful	Possibly Infeasible	Definitely Infeasible	Very Important	Important	Slightly Important	Unimportant
1	Establish contingency plans for priorities in allocating resources														
2	Modify current public health policies to increase the possibility of controlling infectious diseases after nuclear attack														
	n....														

This questionnaire would, of course, be accompanied by written arguments, pro and con, for each proposal listed.

"If the results of this appraisal indicate that an item ranks no higher than 'doubtful' in any category, it would be eliminated from further consideration.

"For the remaining items, some of which would obviously be controversial in one or more aspects, more exacting standards of acceptability would need to be set. The next questionnaire would explore the reasons for any divergence of opinions; it might take this form:

Questionnaire 3

The following items out of the list previously submitted to you have been eliminated for the reasons checked:

Item	Description	Reason for Elimination		
		Undesirable	Infeasible	Unimportant
1	.. .. ..	X		
3	.. ..		X	
4	.. .. ..	X		X
—				
—				
—				

The following items have been accepted as being desirable, feasible, and important.

Item	Description
11	.. .. ..
17	.. .. ..
—	
—	
—	

The remaining items are controversial in one or more respects. In those cases where a check mark is circled, your previously expressed opinion was at variance with the opinions of several of the other respondents. For each, please indicate very briefly why you hold this particular opinion. (For example, if, in Item 6, a check mark in the Desirability column is circled, please explain why you gave Item 6 the desirability rating you did in response to Questionnaire 2.) Alternatively, if on reconsideration you do not feel strongly enough about your previously expressed opinion to defend it, please indicate this by stating a revised rating.

Item	Description	Controversial as to Desirability	Feasibility	Importance	Reason Previous Rating or Revised Rating
2	.. .. ..				
5	.. .. ..				
6	.. .. ..				
–					
–					
–					

> "If the replies to this questionnaire continue to move toward a consensus on some of the proposals, or if for some reason the apparently irreconcilable differences of opinion seem inadequately documented, one or more additional questionnaires may be worthwhile. In form, these would resemble Questionnaire 3.
>
> "What might the final result tell us that we did not already know or could not obtain from less unconventional types of analysis? The answer can be very brief. Many aspects of the postattack recovery problem cannot be handled by standard cost-effectiveness techniques. For example, how can one assess the effect on the arms race of a prewar measure such as the storage of materials for the recovery period? Out example suggests that the Delphi technique offers, at the very least, a way to approach such questions." (Reference 3).

2.7.6 Basic Considerations.

As mentioned previously, whether or not one should use the Delphi procedure depends on the particular situation. However, the analyst, who is considering the procedure, should at least be aware of the following considerations. In contrast to the committee approach where the analyst or initiator's role is essentially complete when the committee has been selected and tasked, the analyst's or initiator's work just begins with the selection of the group in a Delphi exercise. In either case, the selection of the participants is certainly not an easy task. However, in a Delphi exercise, the selection task can be restricted to determining who the experts are. On the other hand, the committee selector must also consider the pers. alities of the participants and the possible face-to-face interaction problems. Certainly, two long-standing enemies could not be placed on the same committee. These same two people, however, could be included in a Delphi exercise.

[3] Quade, Boucher, *Op. Cit.*, pp 334-342.

Perhaps the biggest drawback in applying the procedure at this time is that very few analysts have experience in using the technique. In particular, there is a lack of training in preparing questionnaires and analyzing the results. One should not discount the importance of such training. If the questionnaire is prepared by unqualified people, the answers to the questions may be biased or the questions themselves may not really address the problem. In addition, since the procedure has had limited exposure, it may not be accepted immediately. But this is to be expected with anything new.

Another important consideration in the selection of the Delphi procedure is time, both time available for conducting the analysis and time involved in applying the procedure. Clearly, if there is little time available for developing a consensus of opinion, the Delphi technique may not be a viable alternative. The participants may be spread out geographically and the questionnaire may have to be distributed and collected by mail. Couple this time with processing time and the fact that several iterations will probably be required, and it is not too difficult to imagine the procedure being impractical from a time response standpoint. In addition to the obvious time consideration, if long periods of time elapse between the sessions then the participants may lose their train of thought or may not be able to reproduce their rationale. However, on-line computer systems offer hope for reducing the time involved in a Delphi exercise, and a well planned exercise should take into account the time between questioning sessions. So this need not be a problem.

Next, one should consider whether the group responses can be aggregared meaningfully. If there doesn't exist a meaningful way to aggregate the group responses, one would probably not want to use the Delphi procedure. This potential problem probably would be most prevalent in problem definition and policy formulation applications.

Finally, one must consider whether there are any popular opinions that there may be pressure to conform to. For instance, in the corporate environment it may be known that the president of the company favors a particular policy. This fact may influence the committee chairman to pressure the group or lead the group in the direction of the president's favored policy, even though it may not represent the group's opinion.

While the preceding considerations are not the only things that must be considered, they should, at a minimum, eliminate overlooking the obvious in deciding whether or not to use the Delphi procedure.

2.7.7 Summary.

In situations where one wants to use group judgment to analyze uncertainty, the Delphi procedure provides an alternative to the committee approach in the identification and consolidation activities of a risk analysis. It attempts to improve upon the committee approach by allowing the exchange of information in an environment that reduces the group pressure to conform and removes the impact of the dominant individual. Its primary features are anonymity of the source of information, iteration with controlled feedback, and aggregation of group response.

Experiments give indications that this procedure may provide better results than the committee approach, and that iteration with controlled feedback improves the accuracy of the group response for factual questions. While the evidence is not conclusive, it is at least sufficient to merit consideration of the Delphi procedure in certain problem areas.

2.8 SUMMARY AND CONCLUSIONS

As a means of summarizing the relative attributes of the five techniques described for estimating probability distributions, each of these techniques will be assessed with respect to each of the following criteria:

Ease of Application. Is the technique easy to apply?

Knowledge of Probability Concepts. Must the expert have a good knowledge of probability theory?

Consistency Check. Does the technique offer a means of revealing the consistency of the expert's responses, and thus, a means of validating the resulting probability distribution with respect to the "correctness" of the expert's judgments?

Time required. Is the application of the technique time consuming?

Table 2.21 should provide the analyst with a handy reference table for selecting a technique. It should be emphasized that the problem and the experts background will for the most part determine the set of techniques that one could use. For instance, if there was a short time frame for the study and only one expert was available, we would definitely rule out the Modified Delphi Technique and probably rule out the Modified Churchman-Ackoff Technique. From the remaining techniques, the final selection would probably depend on the expert's probability background. If he had no probability background, we would probably use the standard lottery technique.

TABLE 2.21 ATTRIBUTE SUMMARY

Technique	Ease of Application	Knowledge of Probability	Consistency Check	Time Required
1. Choice Between Gambles (Density Function)	Easy	Must understand the concepts of probability	Not inherent in the technique	Not time consumming
2. Choice Between Gambles (Distribution Function)	Easy	Must understand the concepts of probability	Not inherent in the technique	Not time consumming
3. Standard Lottery	Easy	Need not understand the concept of probability	Not inherent in the technique	Not time consumming
4. Modified Delphi Technique	Not as easy as 1, 2, 3	Must understand the concepts of probability	Not inherent in the technique	Requires a fair amount of time
5. Modified Churchman-Ackoff Technique	Not as easy as 1, 2, 3	Must understand the concepts of probability	Inherent in the technique	Requires a fair amount of time

LITERATURE CITED

1. Hamburg, Morris; Statistical Analysis for Decision Making, Harcourt, Brace & World, Inc., New York, 1970, pp 12-13.

2. Timson, F. J.; Measurement of Technical Performance in Weapon System Development Programs: A Subjective Probability Approach, Memorandum RM-5207-ARPA, December 1968, The Rand Corporation, p 15.

3. Quade, E. S., and Boucher, W. I.; Systems Analysis and Policy Planning, Applications in Defense, Elsevier Publishing Co., New York, 1968, p 334.

4. Dalkey, N.; An Experimental Study of Group Opinion, The Delphi Method, Vol 1, No. 5, September 1969, Futures, p 408.

5. Dalkey, N., Brown, R., and Cochran, S.; Use of Self-Ratings to Improve Group Estimates, Experimental Evaluation of Delphi Procedures, Technological Forecasting 1, 1970, p 283.

BIBLIOGRAPHY

1. Ackoff, R. L., and Churchman, C. W.; An Approximate Measure of Value, Journal of the Operations Research Society of America, 2:172-187, 1954.

2. Barish, Norman N.; Economic Analysis for Engineering and Managerial Decision Making, McGraw-Hill, New York, 1962.

3. Churchman, C. W., Ackoff, R. L., and Arnoff, E. L.; Introduction to Operations Research, John Wiley and Sons, Inc., New York, 1957.

4. Dalkey, N., Brown, R., and Cochran, S.; Use of Self-Ratings to Improve Group Estimates Experimental Evaluation of Delphi Procedures, Technological Forecasting 1, 1970, pp 283-291.

5. Dalkey, N.; Analysis from a Group Opinion Study, Futures Vol 1, No. 6, 1969, pp 541-551.

6. Dalkey, N.; An Experimental Study of Group Opinion, The Delphi Method, Futures Vol 1, No. 5, September 1969, pp 408-426.

7. Definetti, Bruno; Does it Make Sense to Speak of "Good Probability Appraisers?" The Scientist Speculates, I. J. Good, A. J. Mayne, and J. M. Smith (Ed.), Basic Books Inc., New York, 1962, pp 357-364

8. Definetti, Bruno; Recent Suggestions for the Reconciliation of Theories of Probability, Second Berkeley Symposium on Mathematical Statistics and Probability, Jerzy Neyman (Ed.), University of California Press, Berkeley and Los Angeles, California, 1951, pp 217-225.

9. Definetti, Bruno; Foresight: Its Logical Laws, Its Subjective Sources, Trans. Henry E. Kyburg, Jr., Studies in Subjective Probability, Henry E. Kyburg, Jr., and Howard E. Smokler, John Wiley, New York, 1964, pp 93-158.

10. Good, Irving John; The Estimation of Probabilities, An Essay on Modern Bayesian Methods, The MIT Press, Cambridge, 1965.

11. Grayson, C. Jackson, Jr.; Decisions Under Uncertainty: Drilling Decisions by Oil and Gas Operators, Harvard University Division of Research, Graduate School of Business Administration, 1960.

12. Hamburg, Morris; Statistical Analysis for Decision Making, Harcourt, Brace and World, Inc., New York, 1970.

13. Helmer, Olaf; The Application of Cost Effectiveness to Non-Military Government Problems, Systems, Organization, Analysis, Management: A Book of Readings, Ed. by David I. Cleland and William R. King, McGraw-Hill, New York, pp 265-271.

14. Kyburg, Henry E., Jr., and Smokler, Howard E.; Studies in Subjective Probability, John Wiley & Sons, Inc., New York, 1964.

15. Luce, R. Duncan, and Suppes, Patrick; Preference, Utility and Subjective Probability, Handbook of Mathematical Psychology, R. Duncan Luce, Robert R. Bush, and Eugene Galanter (Eds.), Vol III, Chapter 19, John Wiley & Sons, Inc., New York, 1965, pp 249-410.

16. Marschak, Jacob; Actual Vs Consistent Decision Behavior, Behavioral Science, 9 (1964), pp 102-110.

17. Quade, E. S., and Boucher, W. I.; Systems Analysis and Policy Planning, Applications in Defense, Elsevier Publishing Company, New York, 1968.

18. Roberts, Harry V.; Probabilistic Prediction, Journal of the American Statistical Association, 60 (1965), pp 50-62.

19. Sanders, Frederick; On Subjective Probability Forecasting, Journal of Applied Meteorology, 2 (1963), pp 191-201.

20. Savage, Leonard J.; The Foundation of Statistics, Wiley, New York, 1954.

21. Suppes, Patrick; The Role of Subjective Probability and Utility in Decision Making, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol 5: Contributions to Econometrics, Industrial Research, and Psychometry, Jerzy Neyman (Ed.), University of California Press, Berkeley, California, 1956.

22. Timson, F. J.; Measurement of Technical Performance in Weapon System Development Programs: A Subjective Probability Approach, Memorandum RM-5207-ARPA, The Rand Corporation, December 1968.

23. Toda, Majanao; Measurement of Subjective Probability Distributions, Electronic Systems Division Technical Documentary Report No. ESD-TDR-63-407, July 1963.

24. Winkler, Robert L.; The Consensus of Subjective Probability Distributions, Management Science, 15 (1968), pp 61-75.

25. Winkler, Robert L.; The Assessment of Prior Distributions in Bayesian Analysis, Journal of the Americal Statistical Association, 62, (1967), pp 776-800.

Next page is blank.

CHAPTER 3

MONTE CARLO METHODS

3.1 INTRODUCTION

The Monte Carlo method is defined as "Any procedure that involves statistical sampling techniques in obtaining a probabilistic approximation to the solution of a mathematical or physical problem." (Reference 1) The first documented use of the technique appeared in Buffon's treatise "Essai d' Arithmetique Morale" in the year 1777, (Reference 2). In this work, Buffon used the technique to approximate the quantity $1/\pi$ (the ratio of the diameter to the circumference of a circle). In order to approximate this quantity, Buffon created an experiment for which the theoretical probability of occurrence of a particular event was $1/\pi$. He then simulated a series of trials of the experiment and used the fraction of the time the desired event occurred as his estimate of $1/\pi$. This example is typical of many of the early applications where mathematicians were primarily interested in estimating areas or other geometric quantities.

The technique was soon discovered to have numerous applications in many of the applied sciences, particularly Physics, Chemistry and Operations Research. It has also been used extensively in the military's war game modeling efforts. Since these applications are so extensive, no attempt will be made to expound on them in this chapter.

Recently, Monte Carlo has been introduced in risk analysis as a technique for approximating the distribution of critical decision variables such as system time, system cost, and system performance. Suppose, for example, that one of the critical performance characteristics, w, of a weapon system must be greater than or equal to M to meet the postulated threat. Further, w is a function of three independent random variables X, Y, and Z. This is the type situation that could arise in the following example. Consider a project to develop an aircraft and suppose that the performance characteristics that are critical to the aircraft's mission capability are speed, altitude, range, and endurance. For the purposes of this illustration the discussion will be limited to speed. The performance characteristic, speed, is a function of several subsystem and component characteristics. The functional relationship of

---

[1] James, Glenn and James, Robert; Mathematics Dictionary, Van Nostrand Company, Princeton, New Jersey, 1964, p 260.

[2] Buslenko, Nikolai; The Monte Carlo Method, Pergamon Press, Oxford England, 1966, p 4.

these components or subsystem characteristics to the performance characteristic is represented by a design equation. "The design equation for the maximum, constant altitude, level-flight speed of an aircraft is given by

$$V_{MAX} = \frac{195.5 \left[T_{MAX} + \left(T_{MAX}^2 - 1.274 \frac{C_{D_o} S W_g^{\,2}}{\alpha e b^2} \right)^{1/2} \right]^{1/2}}{S C_{D_o}}$$

where T_{MAX} = maximum thrust available (assumed independent of speed),

$C_{D_{(o)}}$ = drag coefficient for zero lift,

S = wing area,

W_g = gross weight of the aircraft,

α = altitude density ratio,

e = efficiency factor, and

b = wing span." (Reference 3)

Now for the postulated threat, the requirements might be for a V_{MAX} to be at least V, and it would be useful to know the probability of meeting or exceeding V. However, to estimate this probability, the distribution of V_{MAX} must be estimated. Given the preceding functional form of V_{MAX} and estimates (objective or subjective) of the distributions of T_{MAX}, C_{D_o}, S, b, e, and W_g, it may be difficult if not impossible to derive the distribution of V_{MAX} exactly with standard statistical tools. However, given the design equation and the distributions of the components, the Monte Carlo technique can be used to construct a sample distribution for V_{MAX}.

Risk analysis applications, such as the preceding, are discussed in detail in Section 3.3. They are followed by a discussion of both the limitations of the technique and the abuses often encountered in practice.

[3] Timson, F. S.; Measurement of Technical Performance Weapon System Development Programs: A Subjective Probability Approach, Memorandum RM-52-7-ARDA, p 12.

Before proceeding with the risk analysis applications, however, a detailed description of the Monte Carlo procedure is provided in Section 3.2.

3.2 DESCRIPTION

The Monte Carlo procedure discussed in this compendium is sometimes referred to as "crude" Monte Carlo. Other sampling procedures exist and can be useful in problems where estimators are computed from the sample. For these problems, the other procedures -- called variance reduction procedures or Monte Carlo techniques -- can be used to get estimators with smaller variance than the "crude" Monte Carlo estimator for the same sample size. These techniques will not be discussed in this compendium. For the reader interested in learning about variance reduction, Reference 4 is recommended.

Basically, the Monte Carlo procedure is used to generate a sample distribution of a random variable whose distribution is unknown by taking into consideration an existing functional relationship between this random variable and others whose distributions are either known or can be estimated (objectively or subjectively). Consider, for example, a random variable Z whose distribution is unknown, but which is linearly related to two independent* random variables X_1 and X_2, whose distributions are known. Suppose the relationship is $Z = 2X_1 + X_2$. Sample values for either X_1 and X_2 are generated by randomly selecting a number between zero and one using a table of uniform random numbers or some computer housed random number generation routine. Given this random number, say ℓ, the sample value is obtained by finding the value such that the probability that the variable of interest does not exceed this value is equal to ℓ. This sample value is unique and can usually be obtained directly from the cumulative distribution function of the variate of interest.

The sample values for X_1 and X_2, generated in this manner, can now be used in conjunction with the relationship $Z = 2X_1 + X_2$ to obtain the corresponding sample value for Z. By repeating this procedure many times, a sample distribution for Z is obtained which approximates the

*Independence is not a necessary condition, but the procedure must be altered slightly to handle dependent variables. A detailed example of a Monte Carlo application for the dependent case is contained in Section 3.3.

[4] Hillier, Frederick S and Lieberman, Gerald J.; Introduction to Operations Research, Holden Day, Inc., San Francisco, California, 1969, pp 452-462.

distribution of Z. Generally, a large number of samples must be taken in order that the distribution for each of the variables is adequately represented in determining the distribution of Z. The actual number of samples required is not easily determined and depends on the specific application.

The following is a suggested approach for determing an adequate sample size for a particular problem. Initially take a sample of size 50. Add another 50 sample values. Compute the sample mean and sample variance and construct a frequency histogram for the initial 50 sample values and for the combined sample of size 100. If the sample means, sample variances or the shape of the histogram change significantly, it may be advisable to increase the step size to 100. In this case, the additional 100 samples would yield a sample size of 200 and from this sample, the mean, variance and the histogram can be compared with the means, variances, and the histograms for the smaller sample sizes. This procedure should be continued until the sample mean, sample variance, and frequency histogram appear to be converging. Note that the step size is critical in the sampling procedure. When 100, 200, or 300 samples have been taken an increase of 50 samples should be enough to detect changes in the sample distribution. However, if it becomes necessary to take in the order of 1000 samples, an increase of 50 samples may no longer be sufficient to detect differences. The possibility exists that steps of 50 will yield successive histograms which are not significantly different and yet the sample distribution has not converged. Thus, it may be necessary at points in the sampling process to increase the step size.

To illustrate how to determine the appropriate sample size, consider a sample from a standard normal distribution. The object is to test the sample distribution of Z for convergence. Each of the component distributions could be tested for convergence but this would not be sufficient to imply the convergence of the distribution for Z. In most cases, the mean and variance of Z are not known and therefore it is necessary to look at successive differences in sample mean and variances to determine convergence. For this example, it is known that the mean is zero and the variance is one.

The results of an initial sample of size 50 from a standard normal distribution are exhibited in Table 3.1. The sample mean is 0.030 and the sample variance is 0.808. By taking an additional 50 samples, the new mean is 0.088 and the variance is 0.875. Noting the differences in the variance and the histograms, one might decide to increase the step size to 100. Continuing with this procedure, it was discovered that 2000 samples are necessary for the histogram and the variance to converge. After 2000 samples, the sample mean is 0.0093 and the sample variance is 0.963. Figure 3.1 shows a sequence of histograms for samples of size 50, 100, 200, 300, 400, 500, and 2000. Note that for 50 and 100 the sample distribution is not shaped like a normal but at 200 it begins to smooth out. After 2000 samples, the histogram looks very much like a normal distribution and the sample mean and variance are close to the values zero and one, respectively.

TABLE 3.1 SAMPLES FROM A NORMAL DISTRIBUTION

SAMPLE MEAN = 0.02999 SAMPLE VARIANCE = 0.80803

SAMPLE VALUES

-1.29421	-1.49644
0.38794	-1.52057
-0.20092	-0.00237
0.89694	-1.10802
0.13029	-1.16548
-0.55059	-0.51442
0.32322	0.68557
1.48328	0.38922
0.00667	-1.27034
1.11237	0.91550
-0.67990	-0.35313
1.56721	-0.33873
-0.17174	0.30416
-0.58121	-0.48749
0.52735	0.90747
0.19513	1.18899
-1.52726	0.53870
0.58078	1.03950
0.57217	-0.07838
0.45787	0.80814
-0.53651	-0.68425
-0.73362	0.69994
0.45915	0.80145
0.70783	1.68601
-1.95595	-0.60831

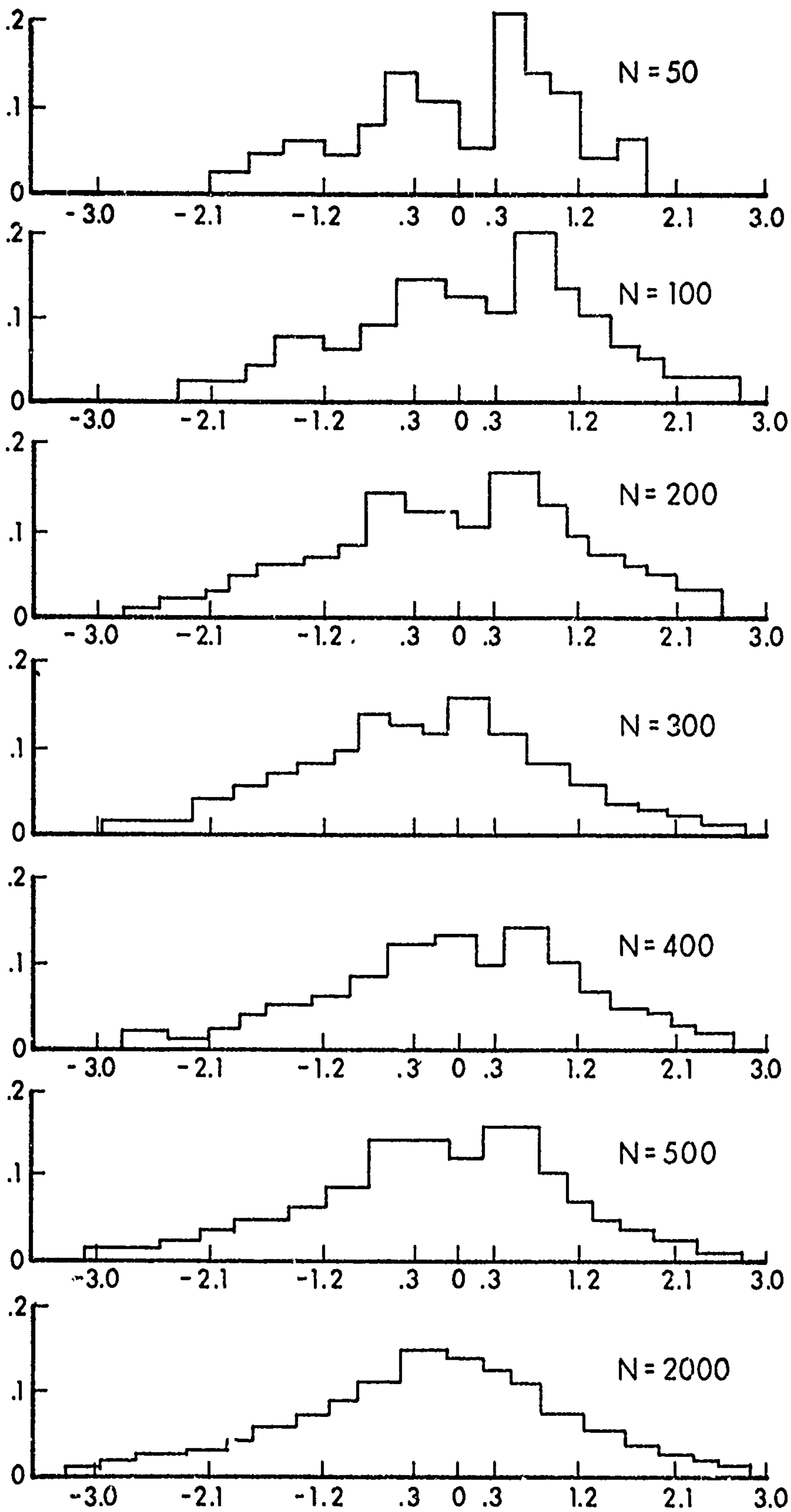


Figure 3.1 Frequency Histograms.

Once an adequate number of samples for Z have been taken, the frequency histogram gives information about the distribution of Z that may aid in the solution to a problem. It indicates whether the distribution is symmetric about the mean or skewed right or left. It also indicates where the modes and the median are.

3.3 RISK ANALYSIS APPLICATIONS

In this section, four examples of Monte Carlo applications to Risk Analysis problems are presented. Two of these examples deal with independent random variables and the others (Examples 2 and 4) illustrate how Monte Carlo can be applied in a dependent situation.

Example 1: The total cost of a system is seldom known with certainty since the estimates are generally based on uncertain contract costs and generally no field experience exists to use in estimating operating cost. However, if the costs are broken down into the elements of Research and Development, Investment Non-Recurring, Investment Recurring and Operating costs, then the uncertainty in these elements can be estimated. In this case the total cost is the sum of the costs associated with each of the individual cost categories. Therefore, given the estimates of the distributions of the elements and assuming that these distributions are independent, it is possible to use Monte Carlo to estimate the distribution of total cost of a system.

Example 2: Consider the problem of comparing the effectiveness of two weapon systems, say System A and System B. Suppose that System A distributes it's payload uniformly in a rectangle centered at the aimpoint and System B distributes its payload uniformly in a circle centered at it's aimpoint. It will also be assumed that wind and other extraneous factors cause the center of the payload patterns to be randomly offset from the aimpoint (say at some random distance and random angle). Given a target shaped as in Figure 3.2, an aimpoint, and a random offset, one can determine the probability that a munition will hit the target. This probability will be the fraction of the payload pattern that intersects the target. Figure 3.3 exhibits a hypothetical situation where the area of the shaded regions divided by the area of the appropriate payload pattern represents the fraction of the bombs that hit the target. Given an aim point, this probability will vary depending on the value of the random offset.

If the distributions of the offset are known (possibly circular normal) for both weapon systems, Monte Carlo can be used to determine a sample value for the offset which will in turn immediately specify the shaded regions. The areas of these shaded regions can then be approximated by repetitive randomly selecting points inside the rectangular pattern and similarly inside the circular pattern and using as the estimate of the area of the shaded region the fraction of the total samples from the respective pattern which fall inside the target area. Thus,

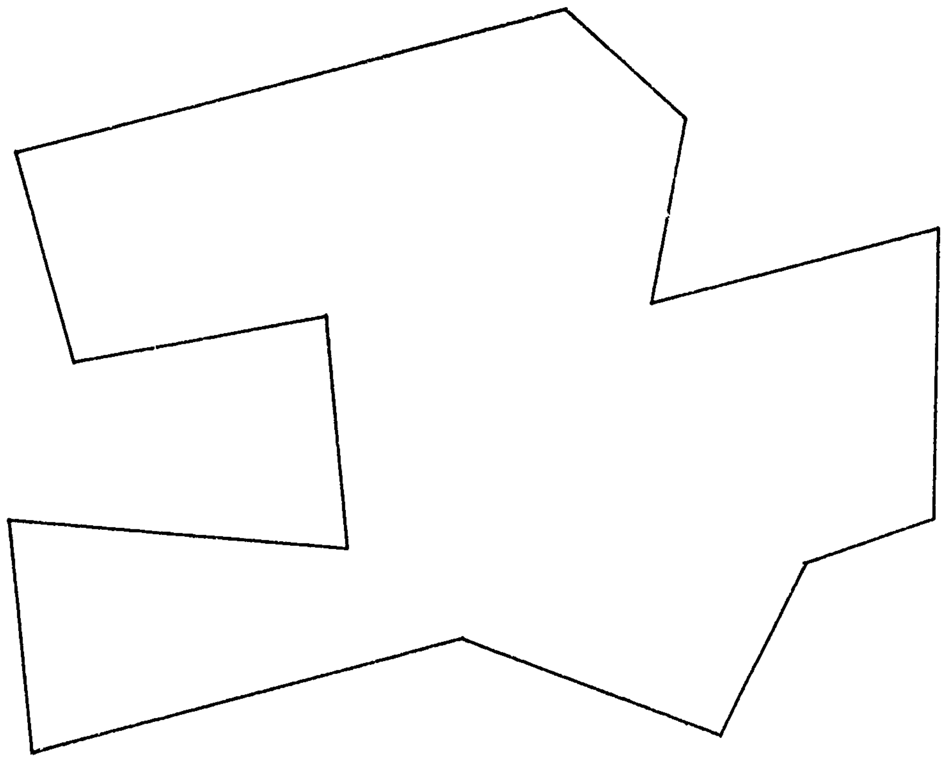


Figure 3.2 Target Region.

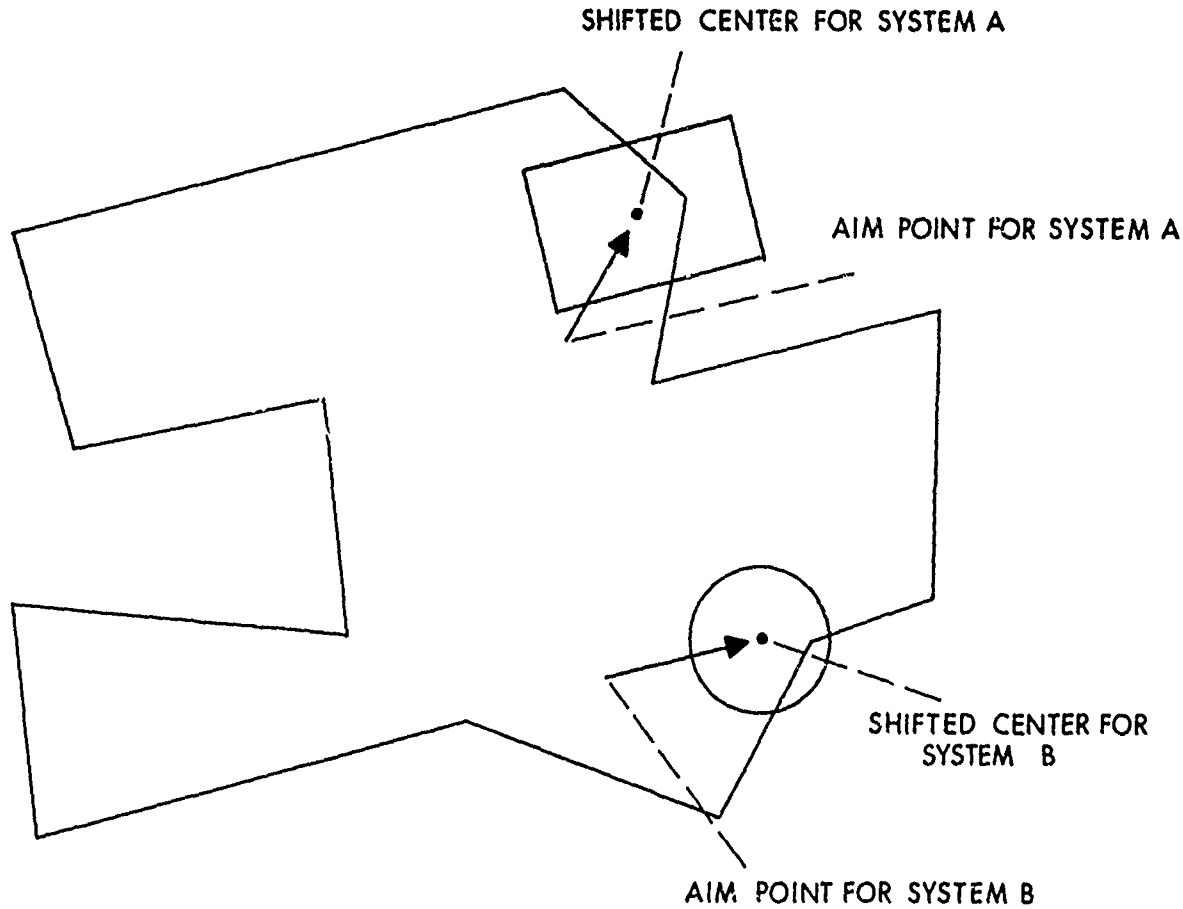


Figure 3.3 Target Region With Payload Pattern Distribution.

given an aim point and a sample value of offset, the probability of hitting the target by a particular weapon system can be obtained using Monte Carlo procedures. Repeating this process by choosing other sample offsets, a sample distribution of the probability of hitting the target can be obtained.

Example 3: "Consider the problem of generating a probability distribution for the life of a two-component electronic device, based on the known probability distributions for the operating lives of the components (tubes), and the design equation which reflects the behavior of the device in terms of the behavior of its components. The design equation in this case states that the device fails if either one of its components fails.

"Suppose that the life characteristics of two electronic tubes are as shown in Figures 13 and 14. To use the Monte Carlo technique, these curves must be changed from density functions to distribution functions. This is done in Figures 15 and 16. Values along the probability scale of the distribution function are selected by means of a table of random numbers. In the case shown, the digits 0 to 9 can be used. (In cases where the probabilities involve two digits, pairs of digits must be selected from a random number table.) When a random number is selected, a horizontal line is drawn from the Y-axis on the cumulative probability curve until it hits one of the vertical lines. This determines a value for the life of the tube. For example, if the random number is 8, then the life of Tube No. 1 is 260 hours. To facilitate working out this example, the data in Figures 15 and 16 are converted into tabular form in Tables 12 and 13.

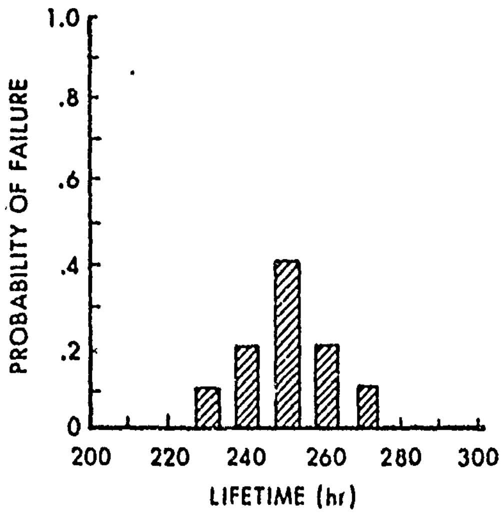


Figure 13-Life Curve for Tube No.1

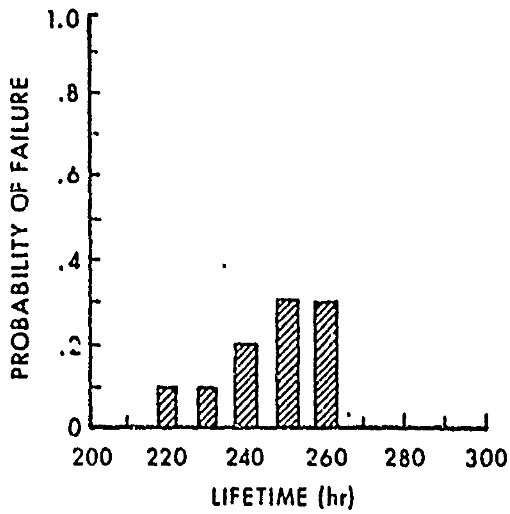


Figure 14-Life Curve for Tube No.2

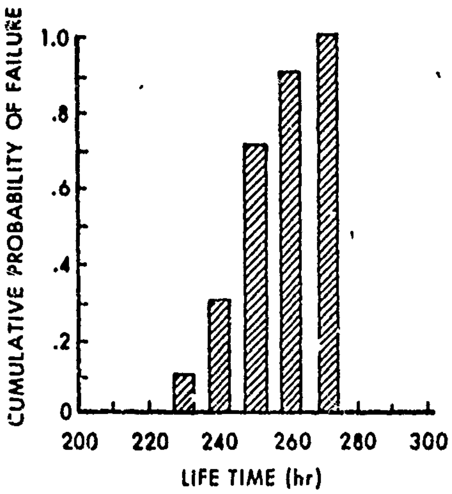


Figure 15-Cumulative Probability of Failure for Tube No.1

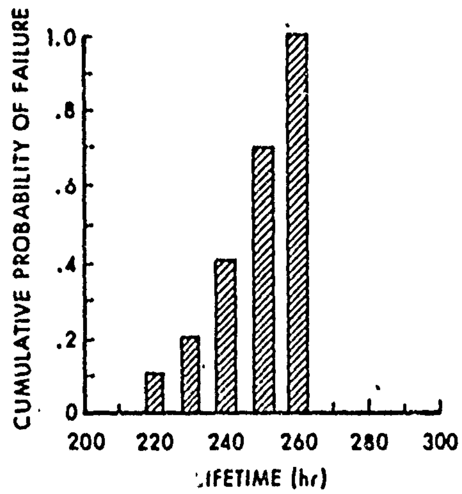


Figure 16-Cumulative Probability of Failure for Tube No.2

"To determine a sample lifetime for the electronic device, a sample lifetime is determined for each tube and the shortest is the lifetime for the device. Suppose the sample lifetime for Tube No. 1 is 260 hours,

TABLE 12

SAMPLING DISTRIBUTION FOR TUBE NO. 1

Random Number	Corresponding Life of Tube No. 1
9	270 hr
8-7	260
6-3	250
2-1	240
0	230

TABLE 13

SAMPLING DISTRIBUTION FOR TUBE NO. 2

Random Number	Corresponding Life of Tube No. 2
9-7	260 hr
6-4	250
3-2	240
1	230
0	220

as determined above. Another random number is drawn for Tube No. 2. Suppose it is 3; the lifetime of Tube No. 2 is then 240 hours. In this instance, the lifetime of the device is 240 hours.

"To determine the distribution of lifetimes for the device, the above process is repeated a large number of times. The frequencies of the observed lifetimes of the device are plotted as a bar chart. This results in a life curve for the device. The number of times that this process is repeated depends on the desired accuracy. The larger the number, the more accurate the resulting life curve will be. Accuracy requirements vary among applications.

"For illustrative purposes, the life curve of the two-tube electronic device is determined by obtaining a sample of 25 pairs of lifetimes for the two tubes. First, a table of the form of Table 14 is set up.

TABLE 14

TWENTY-FIVE SAMPLES OF THE LIFETIME FOR A TWO-TUBE ELECTRONIC DEVICE

Tube No. 1		Tube No. 2		
Random Number[a]	Lifetime (from Table 12)	Random Number	Lifetime (from Table 13)	Device Lifetime
0	230 hr	9	260 hr	230 hr
5	250	4	250	250
4	250	2	240	240
0	230	1	230	230
8	260	0	220	220
0	230	6	250	230
0	230	6	250	230
2	240	6	250	240
5	250	7	260	250
7	260	9	260	260
5	250	2	240	240

TABLE 14 (Continued)

TWENTY-FIVE SAMPLES OF THE LIFETIME FOR A TWO-TUBE ELECTRONIC DEVICE

Tube No. 1		Tube No. 2		
Random Number[a]	Lifetime (from Table 12)	Random Number	Lifetime (From Table 13)	Device Lifetime
8	260 hr	0	220 hr	220 hr
4	250	5	250	250
6	250	8	260	250
5	250	9	260	250
4	250	8	260	250
1	240	2	240	240
3	250	5	250	250
9	270	1	230	230
8	260	9	260	260
4	250	9	260	250
3	250	3	240	240
1	240	0	220	220
5	250	5	250	250
6	250	0	220	220

[a]Random numbers taken from Table 7-3 of Reference 5.

Then random numbers, 25 of each tube, are obtained from Tables 12 and 13. The lifetime of the device is the lifetime of the first tube to fail. A table of the frequencies of the various lifetimes of the device is constructed (Table 15).

TABLE 15

LIFE CURVE DERIVED FROM TABLE 14

Device Lifetime	Frequency	% or Probability
220	4	.16
230	5	.20
240	5	.20
250	9	.36
260	2	.08
	25	1.00

"These results are plotted as probabilities of failure (the life curve) in Figure 17, and as cumulative probabilities of failures in Figure 18." (Reference 3).

[3]Timson, F. S.; Op. Cit., p 59-64.

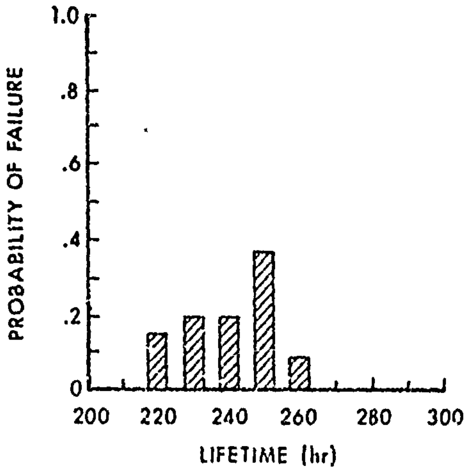


Figure 17-Life Curve for Two-Tube Electronic Device

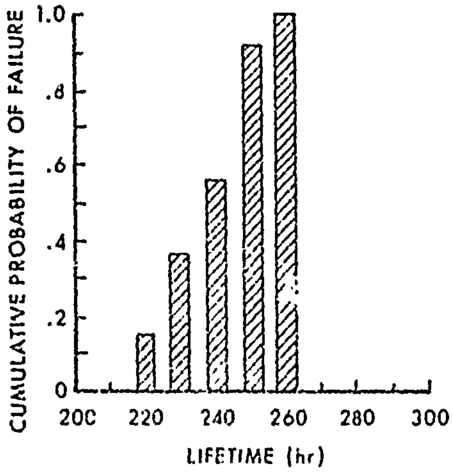
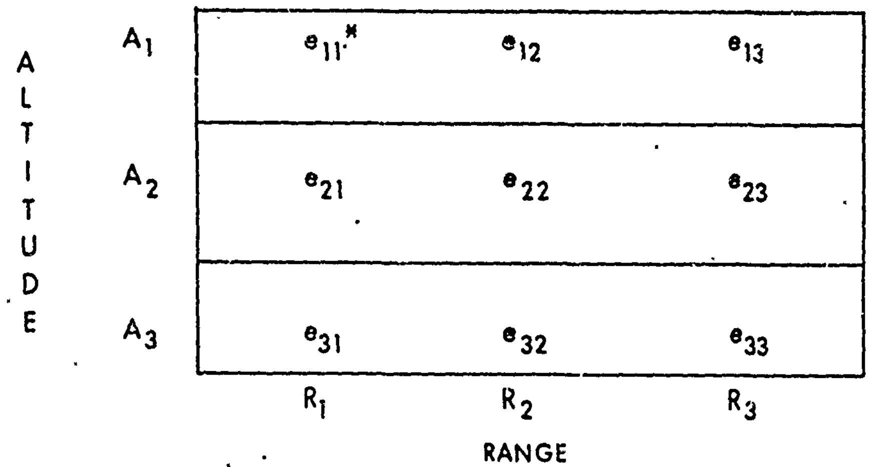


Figure 18-Cumulative Probability of Failure for Two-Tube Electronic Device

Given the information in Figures 17 and 18, it is possible to estimate the probability of exceeding an electronic device lifetime of 250 hours (0.64). This type of information should be of value to the decision maker.

Example 4: Consider the problem of determining the distribution of the effectiveness of a missile against a threat over the entire performance envelope. This performance envelope can be broken down into a discrete range-altitude grid (See Figure 3.4). Assume the effectiveness (e_{ij}) of each range (R; j=1,2,3) and altitude (A_i; i=1,2,3) zone can be estimated, and the frequency of intercept at a particular range given an altitude of attack can be estimated. Then the effectiveness distribution of the missile over the entire performance envelope can be approximated using the Monte Carlo technique. In this instance, the intercept range is dependent on the altitude, but since the distribution of altitude and distributions of range given altitude have been estimated (see Table 3.2), the following Monte Carlo procedure can be used to approximate the distribution of effectiveness.



*e_{11} IS THE EFFECTIVENESS OF THE MISSILE AGAINST THE POSTULATED THREAT AT ALTITUDE A_1 AND RANGE R_1. IN MOST REAL WORLD SITUATIONS e_{IJ} IS A DISTRIBUTION.

Figure 3.4 Range Altitude Grid.

TABLE 3.2 DISTRIBUTION OF ALTITUDE AND DISTRIBUTIONS OF RANGE GIVEN ALTITUDE

Altitude	A_1	A_2	A_3
Distribution	.3	.5	.2

Distribution of Range Given Altitude

Range/Altitude	R_1	R_2	R_3
A_3	.3	.3	.4
A_2	.2	.3	.5
A_1	.8	.1	.1

a. Sample from the distribution of altitude to determine the altitude.

b. Given the altitude, sample from the distribution of range given altitude to determine the sample value of effectiveness.

c. Repeat a and b until an adequate sample distribution has been obtained.

3.4 ADVANTAGES AND DISADVANTAGES

Monte Carlo allows one to approximate the distribution of a random variable when other methods cannot be used. This is often the case when the variable of interest, Z, is a complicated function of several other random variables. In cases where the distribution of Z cannot be derived analytically, provided a large enough sample has been taken, Monte Carlo gives valuable information about the shape of the distribution and also provides a means for obtaining a reasonable estimate of the important population parameters.

Unfortunately, however, in many situations Monte Carlo problems require very large sample sizes which often result in lengthy computer runs. However, since the technique is easy to use and since it always produces results, the length of the computer runs is often overlooked and the technique is used in cases where another approach may be more exact and/or less expensive to implement. For example, suppose we wish to determine the total cost of a system and we know that the component costs are normally distributed and independent. It can be shown that under these conditions the total cost is normally distributed with mean equal to the sum of the component means and with variance equal to the sum of the component variances. For this problem, it would be foolish to use Monte Carlo since it can only give an approximation for a distribution which can be determined analytically.

Likewise, other applications may occur where Monte Carlo is not the best approach. For instance, the variate of interest, Z, may be a relatively simple function of random variables with known distribution. In this case, standard mathematical statistics techniques may result in an exact solution for the distribution of Z.

In other cases, areas which can be approximated by using Monte Carlo can be determined exactly using integral calculus or approximated using a technique such as Simpson's rule using less computer time while maintaining the same degree of accuracy.

Another consideration should be the precision of the technique. Often precision may be inferred when in fact this is not the case. This does not mean that the technique produces imprecise sample distributions, but serves to caution the reader that the output of Monte Carlo is only as good as the input (i.e., estimates and analysis).

There are two specific areas which may degrade the results. First, if the component distributions were derived subjectively, the data may be biased, and secondly, there may be disagreement among the experts. In these cases it is possible to question the expert(s) and test the sensitivity of the distribution of Z to changes in the component distributions.

Another desirable trait of the technique is that it is readily applicable in dependent situations when the conditional distributions can be determined. In some cases, however, a reasonable estimate of the conditional distribution is not available. If the variables are only weakly correlated, the sample distribution generated by sampling independently may be adequate. On the other hand, the following example illustrates how misleading the independence assumption can be when the variables are actually highly correlated. Suppose $Z = X_1 + X_2$ where $X_2 = -X_1$. Z, in this case, is identically zero and hence the variance of Z is also zero. If X_1 has mean zero and variance σ^2, then the Monte Carlo simulation would give a sample distribution for Z with mean approximately zero but with variance approximately equal to $2\sigma^2$. Thus, the analyst must be extremely careful when dependence is suspected. A significant portion of the analysis should be devoted to an examination of the dependence between variables.

3.5 CONCLUSIONS

Monte Carlo is probably the most well known tool available to the risk analyst. As such, it will probably be used widely, but it is hoped that it will be used with discretion. A careful examination of dependence is required in the initial stages of any application.

As indicated in this section, Monte Carlo should have its greatest impact in estimating the uncertainty in the performance capability of a system. It is also of great value in the analysis of probabilistic networks. This application was not discussed since network analysis is discussed in detail elsewhere in this compendium, and in these discussions the role of Monte Carlo is elaborated upon.

It should be emphasized again that to estimate the uncertainty in time and cost other than in probabilistic networks (assuming the component variables are independent) the technique is not usually required. A direct summation procedure can be used just as reasonably with potential computer cost savings.

Next page is blank.

LITERATURE CITED

1. James, Glenn and James, Robert; Mathematics Dictionary, Van Nostrand Company, Princeton, New Jersey, 1964, p 260.

2. Buslenko, Nikolai; The Monte Carlo Method, Pergamon Press, Oxford, England, 1966, p 4.

3. Timson, F. S.; Measurement of Technical Performance in Weapon System Development Programs: A Subjective Probability Approach, Memorandum RM-5207-ARPA, December 1968.

4. Hillier, Frederick S. and Lieberman, Gerald J.; Introduction to Operations Research, Holden Day, Inc., San Francisco, 1969, pp 452-462.

BIBLIOGRAPHY

1. Dienemann, Paul F.; Estimating Cost Uncertainty Using Monte Carlo Techniques, Memorandum RM-5207-ARPA, December 1968.

2. Kahn, H.; Applications of Monte Carlo, Rand Project Report, RM-1237-AEC, 1954.

3. Tocher, K. D.; The Art of Simulation, The English Universities Press Ltd., London, 1963.

4. Van Slyke, R. M.; Operations Research, Volume 11, pp 839-861, 1963.

Next page is blank.

CHAPTER 4

NETWORK ANALYSIS

4.1 INTRODUCTION

4.1.1 Network Concepts.

Before one can discuss the role of network analysis in risk analysis and decision risk analysis, some basic concepts must be introduced. These concepts are a graph, a node, an arc, a network and a path.

The first concept, a graph, is best described by an example. Figure 4.1 is an example of a graph. The circles represent nodes, and the lines joining the nodes are called arcs. Hence a graph is a collection of two or more nodes joined by arcs. Any arc can be charcterized by the pair of nodes that it connects. For exmaple, (1,2) characterizes the arc connecting nodes 1 and 2 in Figure 4.1.

The only difference between a graph and a network is that the arcs in a network have some type of flow in them (See Figure 4.2). One example of a system that can be represented by a network is a development test program. The nodes* in a development test program represent the initiation or completion of various tests, the arcs** represent the actual tests being conducted and the flow in the arcs is time and/or cost involved in testing.

Finally, a path is defined as a sequence of arcs connecting two nodes. For example, the following sequence of arcs form paths between nodes 1 and 4 in Figure 4.2:

PATH (M): (1,2), (2,4)
PATH (N): (1,3), (3,4)
PATH (O): (1,4)

4.1.2 Types of Network Representations.

Given the preceding concepts, it is now possible to describe the different types of network representations and network analysis techniques for analyzing these different types of network representations. The differences in the networks result from assumptions made about the events and the flows in the activities being modeled in the project. As

*Nodes generally refer to events.

**Arcs generally refer to activities or jobs.

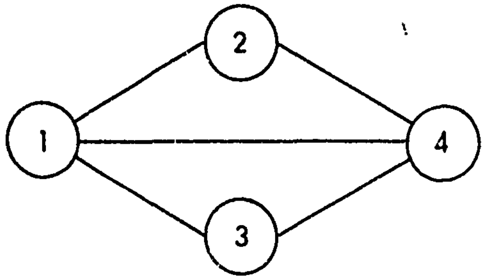


Figure 4.1 Example of a Graph.

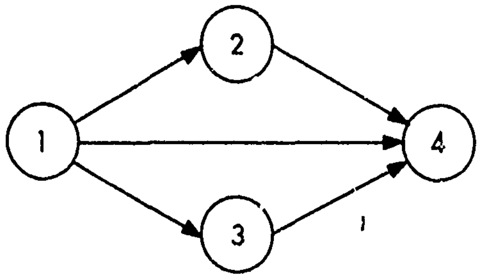
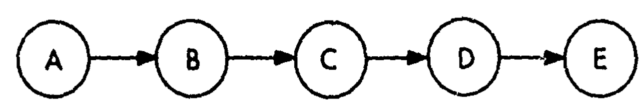


Figure 4.2 Example of a Network.

mentioned previously, for research and development programs the arcs represent activities, the nodes represent events, and the flow in the arcs usually represents time and/or cost.

Three types of network models will be discussed. It should be pointed out that this discussion will be very general and only the major attributes and assumptions about the networks being modeled will be discussed.

The differences in the types of network representations are most easily described by analyzing one example under the varying assumptions about the events and flows in the activities. Changing the oil in a car is the example that will be used throughout this discussion. The first type of possible network representation for describing the events and activities involved in changing the oil in a car is illustrated in Figure 4.3. It is not very realistic for a development program, but it is presented for the sake of completeness.



EVENTS

A - CAR IS LEFT AT THE SERVICE STATION.

B - CAR IS ON THE LIFT.

C - OLD OIL IS DRAINED AND THE FILTER REMOVED.

D - NEW FILTER AND OIL IN THE CAR.

E - CAR OFF THE LIFT.

ACTIVITY TIMES

ARC

AB - 2 MINUTES

BC - 5 MINUTES

CD - 6 MINUTES

DE - 2 MINUTES

Figure 4.3. Changing the Oil in a Car.

Looking at Figure 4.3, there are five milestone events in changing the oil in a car. For this network representation, as well as all others, it is assumed that the events must be completed in a particular sequence in order to complete the project. Certainly, in this example new oil couldn't be added before the old oil is drained. Also, in any R&D project this is a realistic assumption because certain components must be developed before work can be started on others. For instance, if an aircraft is being developed, the type of engine must be developed since different engines will result in different structural requirements for the body. Further, it is assumed that all events must be completed and the completion times are known with certainty (i.e., the events and completion times are assumed to be deterministic).

For R&D programs, these two assumptions are not thought to be very realistic. For rarely are the events and/or activity times known with certainty. For example, there may be several designs under consideration. One may represent an advance in the state of the art while the other may be well within the state of the art. Possibly the more advanced design would be initiated, and the other would be used as a back up position (i.e., in case the primary development was not successful). Based on this example, it is not too difficult to see that the assumption of deterministic events is not very realistic for an R&D project. Similarly, the time may not be known with certainty. Continuing with the same illustration, it may be possible to develop a more advanced design, but how long it will take is uncertain.

Although this kind of network representation is not realistic for R&D projects, it is realistic in the construction industry where tasks for a project are known with certainty. Further, these tasks are repetitive so the assumption of deterministic activity times is also realistic. The Critical Path Method is the name given the network analysis technique developed by DuPont in 1958 to find an efficient method for planning and scheduling the construction of a new facility.

There is a different network representation if, in this hypothetical project, the assumption of deterministic activity times is replaced by the assumption of probabilistic activity times. Probabilistic activity times means that the activity times are not known with certainty (i.e., there exists some distribution of activity times). For instance, the activity time for BC in the oil changing example could vary due to the service station attendant having to pump gas for cars as they arrive at the station. If the attendant is uninterrupted, it may only take 5 minutes to drain the oil and remove the filter. However, this is unlikely since the gasoline station is on the main road, and there is a continual stream of cars stopping for gas. Based on the amount of business this station does, it is possible to estimate the distribution of time to drain the oil and remove the filter as a function of the distribution of interruptions. For instance, the most likely time for completing this activity might be 10 minutes, and the best and worst times might be 5 and 20 minutes respectively. So if this activity time is assumed to be distributed as a triangular distribution with the minimum, most likely and maximum activity times being 5, 10, and 20 minutes respectively, then Figure 4.4 represents the distribution of this activity time. The network analysis technique developed to analyze this type of network representation is Program Evaluation and Review Technique (PERT) PERT was developed in 1958 for planning and controlling the development of the Polaris missile. Even though PERT is more realistic then CPM, it still is deficient for modeling R&D projects since the events are assumed to be deterministic.

Therefore, for modeling R&D programs it is thought that a more realistic network representation would be one that would allow both the events and activity times to be modeled probabilistically. Continuing with the hypothetical example, it may not be possible to replace the oil filter because the type of filter required may not be in stock. In this example, the network would be represented by Figure 4.5. Once the oil is drained from the car, the attendant checks to see if he has the filter in stock. There is a 0.98 probability that he will have it in stock, and a 0.02 probability that the filter is not in stock. In addition, each of the activity times could be modeled probabilistically, but this will not be discussed here since this has already been illustrated for PERT.

There are currently in use or available within the AMC community a number of similar network analyzer programs. All of these can be used to analyze this type of network representation. They are listed here in chronological order of their development:

a. Graphical Evaluation and Review Technique (GERT) developed by Pritsker,

b. Mathematical Network Analyzer (MATHNET) developed by MATHEMATICA,

c. Risk Information System and Cost Analysis (RISCA), a modified version of MATHNET, developed by the Army Logistics Management Center at Fort Lee, Virginia,

d. Statistical Network Analyzer (STATNET), a modified version of MATHNET, developed by Gerald Moeller at the Army Management Engineering Training Agency (AMETA), Rock Island, Illinois, and

e. Venture Evaluation and Review Technique (VERT), a modified version of STATNET, developed by Gerald Moeller of AMETA.

4.1.3 Topics to be Covered.

Obviously, an exhaustive examination of network analyzer techniques is beyond the scope of this compendium. However, two network analyzer techniques, PERT and RISCA, will be discussed in detail in the next two sections.

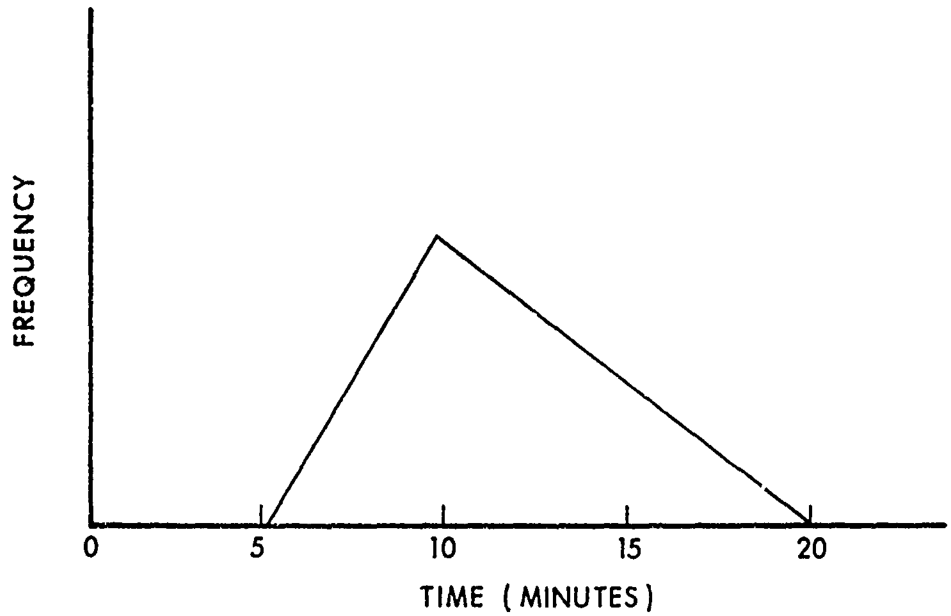
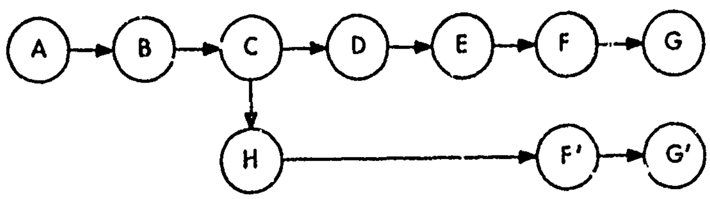


Figure 4.4 Distribution of Activity Time BC.



EVENTS

A – CAR IS LEFT AT THE SERVICE STATION.
B – CAR IS ON THE LIFT.
C – OIL IS DRAINED.
D – FILTER IS REMOVED.
E – NEW FILTER IS IN THE CAR.
F' AND F – NEW OIL IS IN THE CAR.
G' AND G – CAR OFF THE LIFT.
H – FILTER IS NOT REMOVED BECAUSE IT CANNOT BE REPLACED.

Figure 4.5 Changing the Oil in a Car.
(Probabilistic Network)

Recalling the three types of networks, the reason for examining only these two network analyzer techniques is obvious. Since the first network type does not permit one to model either event or activity time and/or cost probabilistically, there is no need to consider the Critical Path method. On the other hand, PERT and RISCA are representative of techniques that can be used to analyze network representations type 2 and 3 respectively.

For both techniques, the following information will be provided:

- Background
- How the technique works
- What type of problems it can handle (emphasis on risk analysis applications)
- Description of output measures
- Benefits and shortcomings
- Example

In addition, there will be a summary and conclusions section that will attempt to examine the role of the type of network analysis provided by PERT and RISCA in risk analysis and decision making within the materiel acquisition process.

4.2 PERT

4.2.1 Background.

The Program Evaluation and Review Technique (PERT) was developed in 1958 for planning and controlling the Polaris Fleet Ballistic Missile Program. It is primarily an R&D management technique used to plan, schedule, and control projects. Because of the success of PERT in the Polaris program, the technique has been applied to other military and commercial projects.

4.2.2 General Discussion.

Before the role of PERT in risk analysis can be defined, the technique will be described as a management technique. The first and probably the most important part of applying any network analysis technique is constructing the network. Network construction consists of graphically describing the sequence of events and activities in the program of interest in node and arc symbols (See Figure 4.6). The reasons that the construction of the network is important are:

a. The results of the analysis are only as valuable as the network representation is realistic.

b. Frequently, this activity forces one to think about event and activity relationships at a level of detail that might not otherwise be done, and as a consequence, a lot of valuable insight may be gained. So in a sense, using PERT can help both structure and discipline the R&D management process.

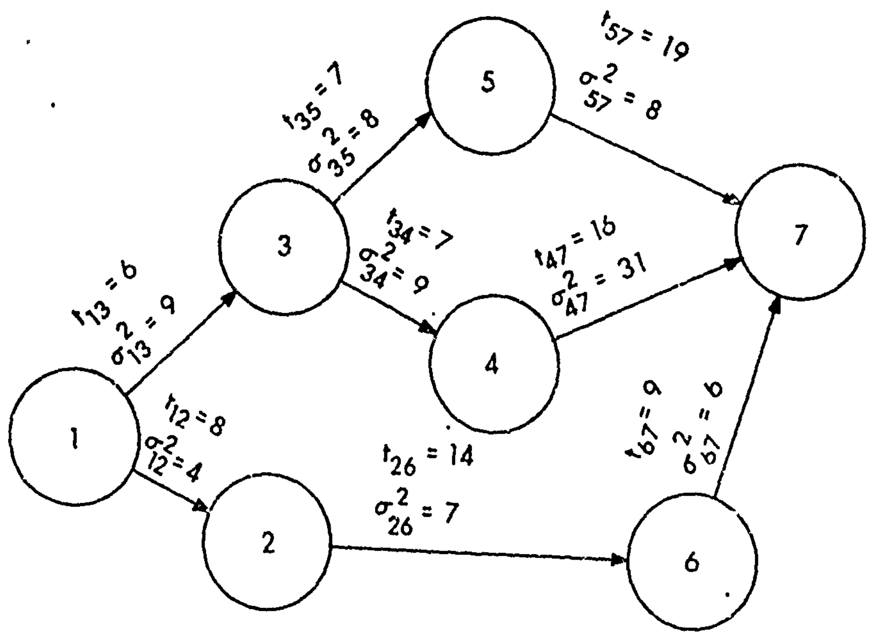


Figure 4.6. Network Representation.

The best way to describe PERT is with a simple example. Looking at Figure 4.6 there are seven events in this hypothetical example. The estimates of the expected activity completion times and the variance of these activity completion times are listed on the appropriate arcs. For example, t_{13} is the expected completion time for activity (13), and σ_{13}^2 is the variance of the activity time of the same activity. These estimates are generally based on subjective estimates of experts and are another crucial element in constructing the network, because the final results will only be as valuable as the estimates are realistic.

At this point, a brief digression is in order to describe the most popular method of obtaining these activity time estimates. What is generally done is:

a. The activity times are assumed to be distributed as a beta distribution (See Figure 4.7.)

b. Individual experts are then interviewed to determine an optimistic (a), pessimistic (b), and most likely (ML) completion time for each activity time.

c. These values are then used to approximate the mean and variance of a beta distribution using the following estimates

$$\text{Mean} = \frac{a + 4ML + b}{6}$$

$$\text{Variance} = \left[\frac{b - a}{6}\right]^2 .$$

A great deal of effort could be expended at this point in discussing the pros and cons of both the selection of the beta distribution and approximations. However, let's just indicate that there is at least one alternative to both the assumption of the beta distribution and the particular estimates. The alternative to the beta is the triangular distribution. The triangular distribution possesses the same desirable distributional proprieties* as the beta distribution; but, unlike the beta distribution, given the optimistic, pessimistic and most likely values a unique triangular distribution is defined. The estimates of the mean and variance of the triangular distribution are:

$$\text{Mean} = \frac{a + ML + b}{3}$$

$$\text{Variance} = \frac{1}{18}\left[(b-a)^2 + (ML-a)(ML-b)\right] .$$

Next, given the network and the estimates of the mean and variance of activity times, the critical path through the network is determined. The critical path is the path of the earliest project completion time (i.e., the longest path). This path is important because any delays in any of the activities on this path will delay the entire project by the same amount of time. So any manager interested in completing the project by a prescribed date should be monitoring the activities on the critical path carefully.

Of course, determining the critical path is not a simple matter for complex projects because enumeration of paths may not be possible. However, there are methods, such as dynamic programming, for handling these types of enumeration problems.

---

*unimodal and can be skewed.

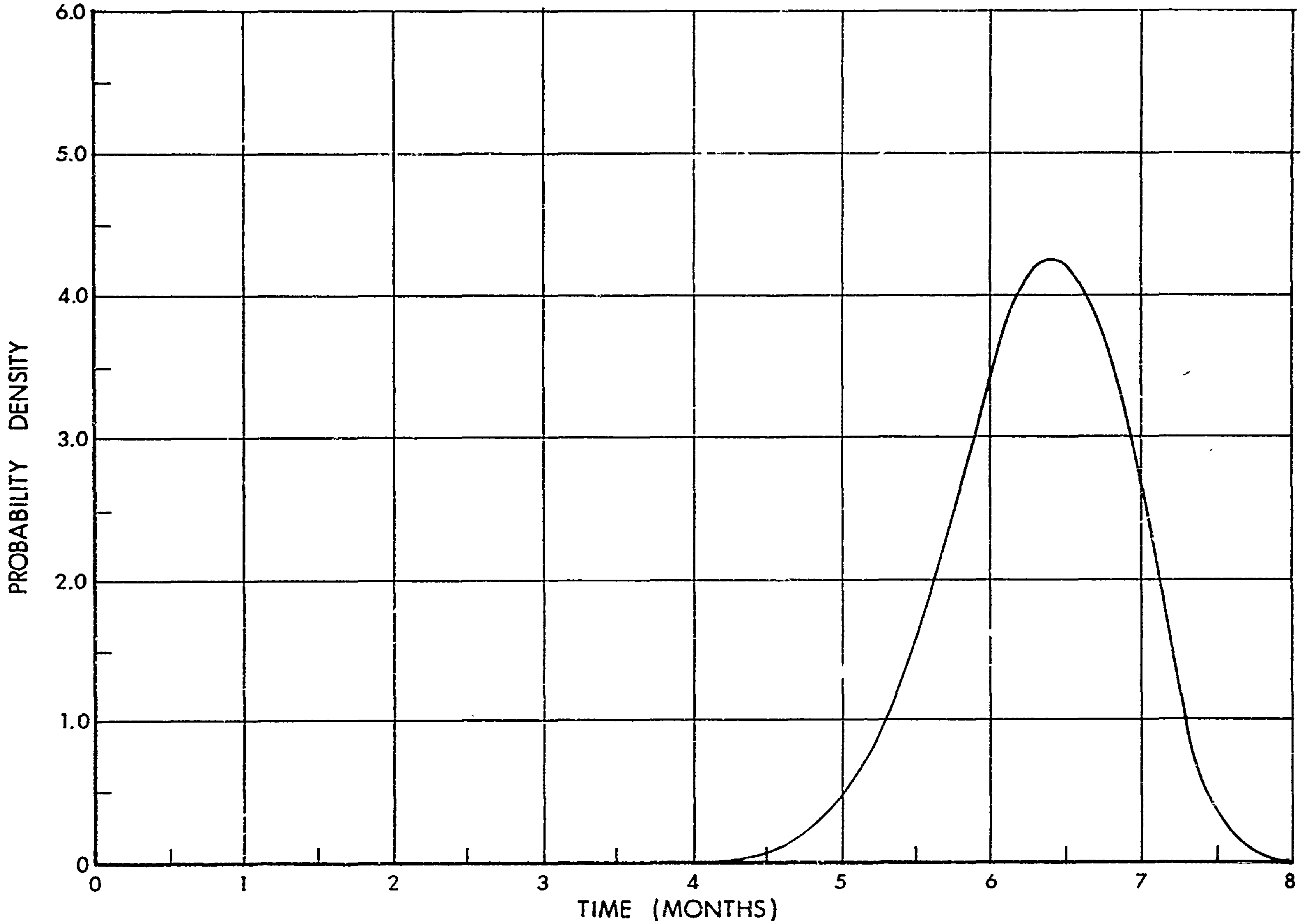


Figure 4.7 Beta Distribution.

One dynamic program algorithm for determining the critical path is the forward pass method.

The nomenclature for subsequent critical path calculations will be the following:

t_{ij} = expected time associated with the directed arc from node it to node j.

f_i = expected time for the longest path from node 1 to node i, i.e., the earliest completion time for event i (also called the "label" of node i).

g_i = the number of the node from which node i was labeled, (i.e., the number of the predecessor to node i on the longest path) for each i = 1,...,N, where N is the terminal node (project completion). The following relations are valid for any PERT network,

$$f_1 = 0$$

$$f_j = \max\,(t_{ij} + f_i) \quad j = 2,\ldots,N.$$ [1]

In the example described by the network in Figure 4.6, the values of f_i and g_i are:

$f_1 = 0$	$g_1 = --$
$f_2 = t_{12}+f_1 = 8+0 = 8$	$g_2 = 1$
$f_3 = t_{13}+f_1 = 6+0 = 6$	$g_3 = 3$
$f_4 = t_{34}+f_3 = 7+6 = 13$	$g_4 = 3$
$f_5 = t_{35}+f_3 = 7+6 = 13$	$g_5 = 3$
$f_6 = t_{26}+f_2 = 14+8 = 22$	$g_6 = 2$
$f_7 = \max\,(t_{57}+f_5, t_{47}+f_4, t_{67}+f_6)$	$g_7 = 5.$
$= \max\,(32,29,31) = 32$	

So the earliest project completion time is 32 months and the longest path is constructed from the g_is as follows:

[1] ALMC, A Course of Instruction in Risk Analysis, ALM-3315-H, United States Army Logistics Management Center, Fort Lee, Virginia, p V-43.

a. the last arc on the path is (j,N) where $g_N = j$.

b. the next to the last arc is (k,j) where $g_j = k$.

c. this construction continues until some $g_s = 1$, then the first arc is (1,s).

The critical path is then (1,s),...,(k,j)(j,N).

Next, PERT analysis provides a method for determining the latest allowable date for completing an event and still keeping the project on schedule. This computation is relative to the time required to complete the project.

"Let h_i be the symbol for the latest allowable completion time for event i. By definition, h_i is the latest time by which an event must be completed to keep the project on schedule. Let T_R represent the required project completion date. The required project completion date might be less than, greater than or equal to the earliest project completion time.

"The values of h_i are computed from each event by working backward from the last event and computing values of h according to the following relationships:

$$h_N = T_R$$

$$h_i = \text{Min } (h_j - t_{ij}) \quad j = N,\ldots,1$$

where N again represents the number of the terminal node or final event. These formulas can be expressed in words as follows:

a. The latest allowable completion date for the project h_N, is identical to the required project completion time, T_R.

b. To compute the value h_i for an event i, subtract the expected value of time for activity (i,j) from the value of latest allowable completion time, h_i, for the successor event, j.

c. If more than one value of h_i is obtained, the smallest value is selected." (Reference 1).

[1] Ibid, pp V-43, V-44.

For the example in Figure 4.6, assuming the project begins at time zero and the time required for project completion is the expected completion time of the critical path (32), the latest allowable completion dates are:

$$h_7 = 32$$

$$h_6 = h_7 - t_{67} = 32-9 = 23$$

$$h_5 = h_7 - t_{57} = 32-19 = 13$$

$$h_4 = h_7 - t_{47} = 32-16 = 16$$

$$h_3 = \text{Min}\ (h_4 - t_{34}, h_5 - t_{35}) = \text{Min}\ (9,6) = 6$$

$$h_2 = h_6 - t_{26} = 23-14 = 9$$

$$h_1 = \text{Min}\ (h_3 - t_{13}, h_2 - t_{12}) = \text{Min}\ (0,1) = 0$$

In this example, h_6 indicates that event 6 must be completed 23 months after the project begins in order to complete the project in 32 months.

Finally, using "the values of the earliest completion time for event i, f_i, and latest allowable completion time, h_i, it is possible to calculate the slack of an event i. The slack of event i is denoted by S_i and is given by:

$$S_i = h_i - f_i$$"[1]

Slack indicates whether there is more than enough time to complete the project by a required date (positive slack), whether there is adequate time to complete the project by a required date (zero slack), or whether there is not enough time to complete project by a required date (negative slack). Continuing with the same example in Figure 4.6 the slack for the events is:

$S_1 = 0$ $\quad$ $S_5 = 0$

$S_2 = 1$ $\quad$ $S_6 = 1$

$S_3 = 0$ $\quad$ $S_7 = 0$.

$S_4 = 3$

[1] Loc. Cit., p V-16.

The way things are currently scheduled, the second event's completion date could slip a month and still not delay the project completion. Note that the slack of all of the events on the critical path is zero. However, this is only true when the required project completion time is equal to the critical path completion time.

Therefore, PERT, as it was originally designed, provides management with information that should enable them to devote the majority of their energies to managing the activities that are most crucial to the timely completion of the project or for scheduling projects for completion by a prescribed date.

In addition, the PERT critical path information is sometimes used to estimate the probability of meeting required completion dates. This is done by making the assumption that the probability of the project being completed by some required date will be approximately equal to the probability that the critical path will be completed by the required date. It is also assumed that the activities and activity times on the critical path are independent. If the number of activities is large, the distribution of the completion times for the critical paths will be approximately normally distributed with mean equal to the sum of the means of activity times on the critical path and the variance equal to the sum of the variances of activity times on the critical path. In this example,* the distribution of completion times would be normal with mean equal to 32 months and standard deviation equal to 5 months. Using this information, the probability of completing the project within 35 months would be 0.72, i.e.,

$$P[X \leq 35] = P[Z \leq \frac{35 - 32}{5}] = P[Z \leq 0.6] = 0.72 \quad .$$

4.2.3 Possible Problems.

Up to this point, the discussion has been for the most part positive, but there are several potential problems. The first potential problem is that PERT network representations may not be adequate for describing events. Recall that in all PERT evaluations the events are assumed to be deterministic. This may not be a bad assumption for short range planning in a development program, but certainly for long range planning there is a strong possibility that the events are uncertain. For example, in the development of a truck, a diesel engine might be preferred. However, development of a diesel engine might represent an advance in the state of the art, and the chances of successful development might be 50/50. The developing agency may plan to initiate the development of the diesel engine, and if it is unsuccessful then an existing gasoline engine will be modified for the truck.

*In a practical situation, three activities would not be large enough.

Clearly, the events in this development are uncertain, but PERT would not allow one to analyze such a situation. It is not meant to imply that PERT is not useful for planning and controlling R&D projects, but only to caution the analyst to PERT's limitations. There is nothing wrong with using PERT to analyze a network where the events are deterministic or can be assumed to be deterministic without any great consequences. However, if the assumption cannot be made, the results may be misleading. The responsibility rests with the analyst to determine if the events can be assumed to be deterministic.

Next, the greatest potential problem in using PERT is that the usual method for determining the critical path may not include critical activities. Recall that PERT uses a dynamic program algorithm to determine the critical path. If P_i, i=1,...,N represents the completion time of the N possible paths, the forward pass method selects the maximum P_i, (i.e., Critical Path = Max $[P_i]$ i = 1,...,N). This method is fine if the activity times are deterministic, but they aren't. What is desired is the expected value of the maximum path, (i.e., $E[\text{Max}\ (P_i)]$). Unfortunately, it is not practical to derive this analytically except in extremely simple cases.

Faced with this problem, there are two possible solutions. One is to make some simplifying assumptions that allow one to obtain an approximate answer, and the other is to use simulation methods. In most applications, the following simplifying assumption is made; "the largest total elapsed time for the project <u>always</u> occurs on the path with the largest <u>expected</u> total elapsed time." (Reference 1).

Using this assumption, the maximum expected value of the paths is determined, i.e., Max $[E(P_i)]$. In general, however, $E[\text{Max}(P_i)] \neq \text{MAX}\ [E(P_i)]$ i = 1,...,N. As a result, PERT estimates of the expected value of project completion times are always low. There is always the chance that another path's project completion time could exceed the expected project completion times. Consider for example path (1,3)(3,4)(4,7). The probability of the completion time for this path exceeding 32 months is*

$$P[X > 32] = 1 - P[X \leq 32] = 1 - P[Z \leq \frac{32 - 29}{7}]$$

$$= 1 - P[Z \leq 0.429] = 1 - 0.66 = 0.34 \ .$$

*Assuming the path completion times are normally distributed.

[1] <u>Loc. Cit.</u>, p V-39.

In this example, the activities on the path may be just as critical as the activities on the critical path. Therefore, the potential problem lies in that all critical activities may not be identified with this method. The net result might be a program delay that results from delays in unidentified critical activities.

What can be done? For one thing, being aware of this potential problem should at least caution the manager not to ignore other activities. In addition, simulation offers a reasonable solution to the problem. The procedure for determining the critical path would then be:

a. Randomly sample from each activity time distribution,

b. Using these values, use the forward pass method to determine the critical path,

c. Record both the completion time for the critical path and the activities which compose the critical path, and

d. Repeat the above steps n times.

In this case the n completion times would serve as the distribution of minimum project completion time, and the activities on the n paths could be placed in histogram form to display the frequency of occurrence of each activity on the critical path.

Of course this is not the solution to all problems either. It may be both difficult and expensive to implement such a simulation for large networks. In addition, the traditional PERT estimates of the beta distribution can not be used in the simulation without first making some assumptions because a unique beta distribution was not defined.

While the preceding potential problem should not discourage any potential user, they should serve as a reasonable set of criterion for interpreting results for planning, scheduling and controlling programs.

4.2.4 Risk Analysis Applications.

So far, no mention of the role of PERT in risk analysis has been made. It should be emphasized that risk analysis should be an integral part of program management. PERT was developed to help R&D program managers cope with uncertainty in planning, scheduling, and controlling activities in their program. However, the emphasis in management is on identifying the greatest potential schedule risks. Identification of risks should reduce the uncertainty in planning and scheduling actiiities, and the manager should be able to more effectively control the program by closely monitoring critical activities. In addition, the risk of not meeting program schedule deadlines can be estimated also. Therefore, PERT is thought to be most applicable as a tool for program managers. Of course, all the preceding comments only apply to situations where PERT can reasonably be applied. It is the responsibility of both the analyst and the manager to make sure that the assumptions made in PERT do not misrepresent reality to the extent that the results of the analysis are suspect.

4.3 RISCA

4.3.1 Background.

RISCA is a modified version of another network analyzer program, MATHNET (Reference 2). It was modified by the Army Logistics Management Center (ALMC) at Fort Lee, Virginia. Under PROMAP 70, the responsibility for instruction in Risk Analysis was assigned to ALMC. In order to expedite the program, a contract was let to MATHEMATICA by the Army Research Office to develop a course of instruction in Risk Analysis. MATHNET was developed as a teaching aid for this course. The materials were then used by ALMC as the foundation for the current course of instruction in risk analysis. ALMC not only has expanded on the course content, but they have also modified MATHNET and called their version of the program RISCA. For a detailed discussion and comparison of MATHNET and RISCA, see Reference 3.

4.3.2 General Description.

RISCA is a computer program that allows one to analyze systems that can be represented by a general class of networks. Since the events and activity times and/or costs can be modeled probabilistically, the analysis is accomplished by simulation. The output consists of a frequency distribution for all possible terminal events and the corresponding time and/or cost distribution for each terminal event. In addition, the distribution of time and/or cost weighted over all possible terminal events is estimated.

Before one can analyze a system with RISCA, the system must be represented by a network. As mentioned previously, many of the benefits derived from analyzing a network result from the analysis and thinking that goes into the construction of the network. Consider, for example, the development of a tank where there are several alternate designs. Describing the sequence of events for alternate development programs for each design should provide valuable insight into the types of problems that one is likely to encounter in each program.

---

[2] Mathematica; MATHNET, Preliminary Edition, August 1970, Princeton, New Jersey.

[3] Brooks, W., Foster, W., and Maruyama, R.; MATHNET and RISCA (Network Analyzer Program) A Users Manual, to be published as an AMSAA Technical Report, 1972, Aberdeen Proving Ground, Maryland.

In order to better explain how RISCA works, the simulation of the oil changing example (Figure 4.5) from the introduction is described.

Since the event "removing the oil filter" is the only uncertain event, the probabilistic event network can be described in in terms of two deterministic event sub-networks. One sub-network represents the events and activities involved in changing the oil and the oil filter, and the other represents the events and activities involved in changing only the oil. In this simple example there is only one path in each deterministic event sub-network, but in more realistic problems there will almost certainly be several possible paths in a sub-network.

Monte Carlo procedures are used to construct a deterministic event sub-network from the probabilistic event network. Each of the sub-networks have a terminal event whose completion time is determined by sampling all the activity completion time distributions in the sub-network using Monte Carlo procedures. All the potential paths in the sub-network are then investigated using these sample activity times. The completion time for terminating the sub-network is the path with the longest completion time. In the oil changing example, if the sub-network constructed was the one involving removal of the oil filter, the activity time distributions in this sub-network would be sampled. Since only one path exists, these sample values would then be summed to estimate the sub-network completion time.

It should be noted that determining the maximum time path is not as straight forward as in PERT analysis because of the network logic permitted in RISCA. This network logic and the perturbations in the maximization procedure it causes will not be discussed here. The interested reader is again referred to Reference 3.

In addition, the cost of all activities in the sub-network would be sampled and summed to estimate completion costs. However, in this example, costs were not considered.

The preceding procedure is repeated many times and the sample distributions of terminal events and time and/or cost are constructed. It should be pointed out that RISCA does not determine a deterministic event sub-network first, but the procedure was so described for the sake of clarity. Actually, the deterministic event sub-network is constructed and the corresponding time and cost estimates are accumulated as the network is analyzed sequentially.

---

[3] Ibid.

For this example, assume 300 iterations have been run. RISCA's output would consist of frequency histograms of the percentage of times each terminal event was selected as shown in Figure 4.8 and the completion time distributions shown in Figures 4.9, 4.10, and 4.11. In addition, a cumulative distribution of time for each completion time distribution in Figures 4.9, 4.10, and 4.11 would be provided. Further discussion and interpretation of the output is deferred until the example output is described.

It should be pointed out that costs are not considered in this example (i.e., the cost of oil and a filter is known with certainty). However if costs were uncertain, there would also be distributions of cost for all terminal events.

In most real world problems, costs are uncertain. For the analyst modeling the network, there are two options available for estimating the activity costs with RISCA: (1) the cost can be estimated independent of time by running a separate simulation or (2) the cost can be estimated as a linear function of time in the same simulation, [i.e., Cost = (fixed cost) + (variable cost x time)].

In the next few pages the applications and benefits to be derived from doing this type of analysis will be discussed.

4.3.3 Risk Analysis and Decision Risk Analysis Applications.

RISCA provides a framework for modeling uncertain events, activity times, and activity costs in a development program and simulating the corresponding network representations. In the context of risk analysis, RISCA provides a method for quantifying development program time and cost risks in a meaningful, summarized form. For example, consider a tank being developed to replace an existing system by 1980. What are the chances of being fully deployed by 1980, or what is the risk of not being fully deployed by 1980? Given a network representation that realistically lays out the sequence of events and activities leading up to full deployment and estimates of the uncertainty in the activity completion times and events, a network analyzer program, like RISCA, could be used to estimate the chances of meeting or not meeting this deployment date.

Several comments should be made at this point. First of all, the estimate of risk is only as good as the analysis that went into modeling the network. If a marginal effort is devoted to structuring the network, the results may not be worthwhile, or they may be misleading. In addition to the obvious reason for carefully structuring the network, there is another more subtle reason that this modeling effort should be emphasized. Sturcturing a network will probably force the user to think about the interaction and consequences of events at a level of detail that he would probably not do otherwise. Frequently, such a probing examination gives the user additional insight and perspective into potential problems.

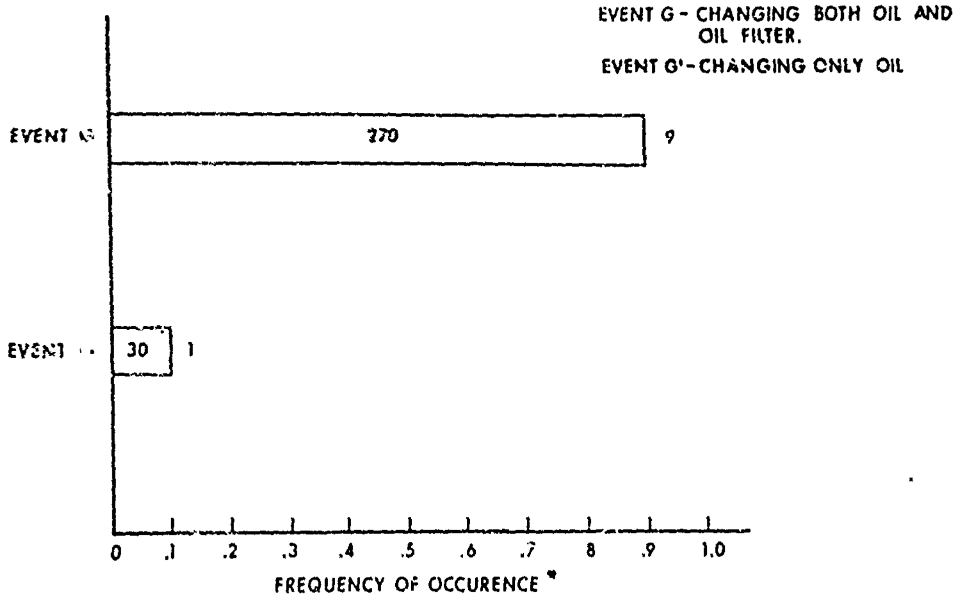


Figure 4.8. Possible Terminal Events.

* PERCENT OF TIMES THE TERMINAL EVENT WAS SELECTED.

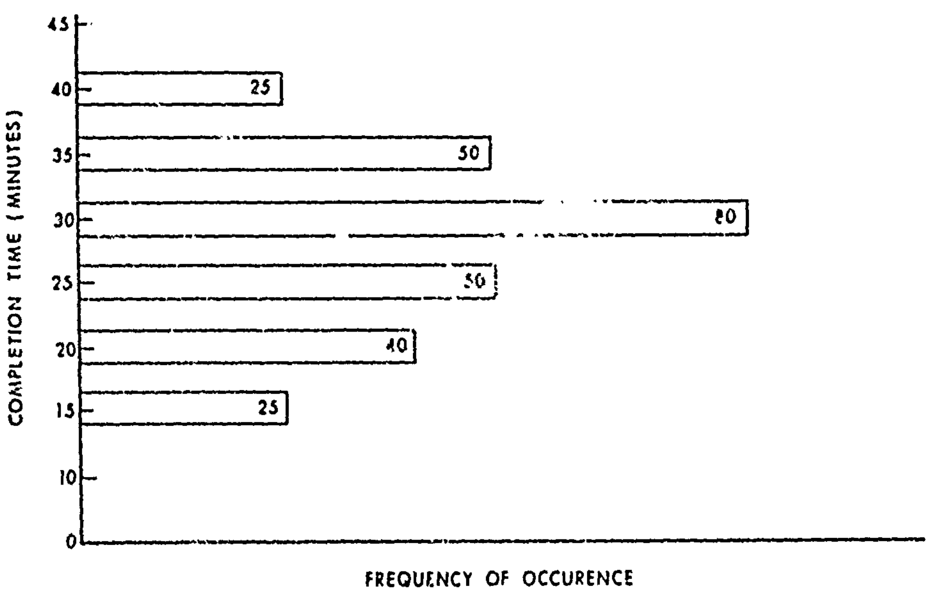


Figure 4.9. Frequency Histogram of Completion Time for Terminal Event G.

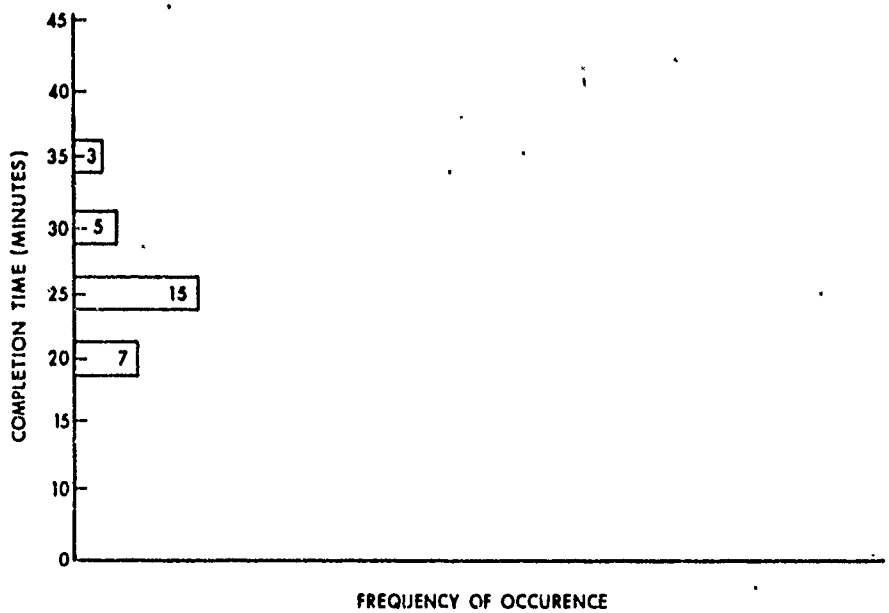


Figure 4.10. Frequency Histogram of Completion Times for Terminal Event G'.

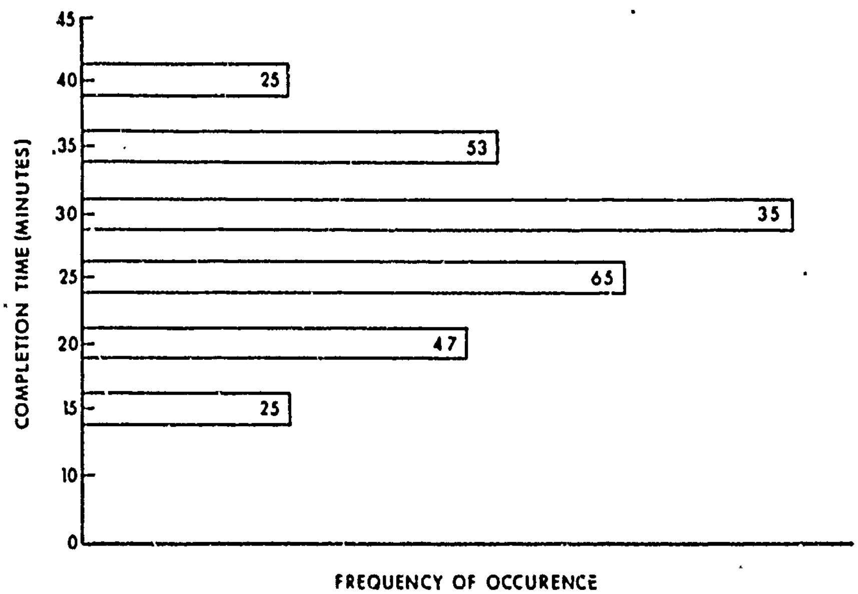


Figure 4.11. Frequency Histogram for Project Completion Times Weighted Over Terminal Events G and G'.

Next, the risks will change over time. This results from the fact that as time passes more information is gained. Two things can happen. One is that the outcome of previously uncertain events may now be known or the events leading up to the uncertain events may indicate that the event is more or less uncertain than it was previously thought to be. The other is that there exists unknown uncertainty, and as time passes the unknown uncertainties may surface. For instance, a problem may be surfaced during test and evaluation that could not be predicted during design. In this case, there must be a continuing or periodic update of the network representation in order for the risk estimates to be valid.

In addition, network analysis can be a valuable tool in a decision risk analysis. Network analysis can provide the foundation of the structure for an evaluation for decision-making purposes. For example, the decision might be to choose between alternative systems for meeting a particular set of requirements. For this type of problem, a network could be structured for each alternative. Each of the networks could then be simulated, and the chances of successfully developing each alternative system and the corresponding development costs and completion times could be estimated. Given this information, the analyst has several relative measures for comparing the alternatives. In addition to the estimates of time and cost, the analyst has the information in a form that will allow him to make relative time and cost risk statements for each alternative, if there exist time and cost constraints. Of course, there is more to doing a decision risk analysis than network analysis, but this type of analysis can be a valuable asset.

In the next section, a hypothetical problem will be described, structured as a network, and simulated using RISCA. The output will then be used to make risk statements and a decision. The emphasis in this example will be on structuring the problem as a network. The only piece of information not given is how to input this network representation into the program. The interested reader is again referred to Reference 3.

4.3.4 Example Problem.

During the past few years a certain worker has noted that occasionally he has been late for work. However, because of his industriousness, the worker has never feared reprimand. Never, that is, until the management adopted a new tardiness policy. The policy dictates that the number of minutes each employee is late during the succeeding 100 weeks (500 working days) will be cumulated and deducted from the final pay check if the employee is late more than 10 percent of the time.

[3] Loc. Cit.

Although the worker knows that he has been late for work, he does not know if he has been late over 10 percent of the time. Because this policy could potentially cost the worker quite a bit of money, he has decided to carefully evaluate his routes to work and determine the chances (risk) of his being affected by this new policy.

Figure 4.12 provides a graphic illustration of the workers's route to work. The following is a description of the potential hazards, delays, and decision points.

a. A Fork in the Road. Not far from the workers's house is a fork in the road. Both roads ultimately lead to work. One, however, is considered an alternate route because of its greater length. On occasion, this longer route must be taken because of intermitten construction along the primary route.

b. Fuel Problems. The worker is assured of not having to to stop for gas if the primary route is taken. However, if the alternate route must be taken, there is a chance that a 5-minute fuel stop must be made.

c. Bumpy Road. Beyond the gas station is a stretch of extremely rough road. The worker has, on occasion, been forced to make a time consuming stop to change a flat tire along this segment of the route.

d. Rider Stop. The next milestone along the route is the house of a fellow worker. The man in our example must stop here every day to give him a ride. Most of the time, the rider is waiting outside and no appreciable time is consumed in the pick-up. However, if the rider is not waiting, our worker must make a five minute stop to check. The probability that the rider is not waiting outside is a function of the delays along the route.

Using standard network notation (i.e., arcs representing activities and nodes representing events or decision points) the route to work can be represented as illustrated in Figure 4.13. In Figure 4.13, the general flow of network is from left to right. The nodes which involve chance (fork in the road, stop for gas, bumpy road, and check rider) are characterized by an arc entering from the left and two arcs exiting. The arc which is chosen to exit these chance nodes is determined probabilistically.

The remaining nodes identify other significant events in the network and mark the beginning and end of segments along the route. The arcs connecting all nodes contain distribution and time data inherent in the system. These arcs are also equipped to handle cost data which is not considered in this example.

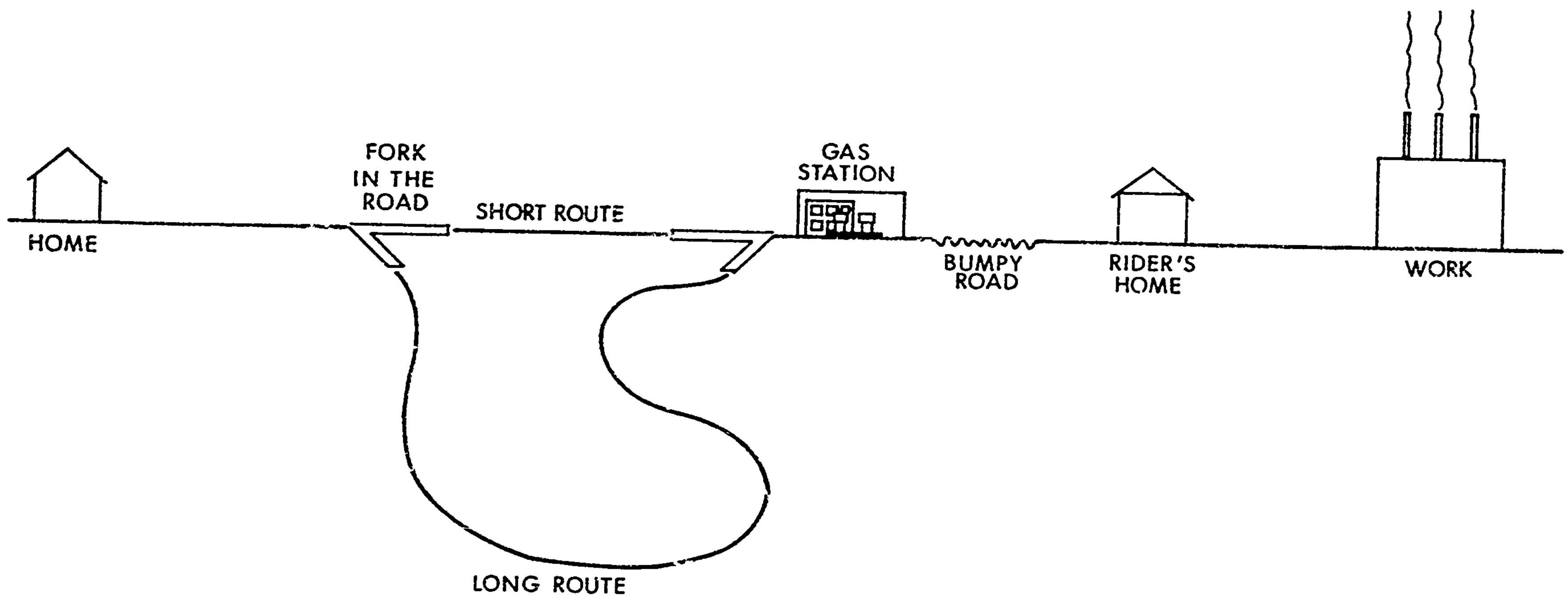


Figure 4.12 Graphical Represr ntation of the Driving to Work Problem.

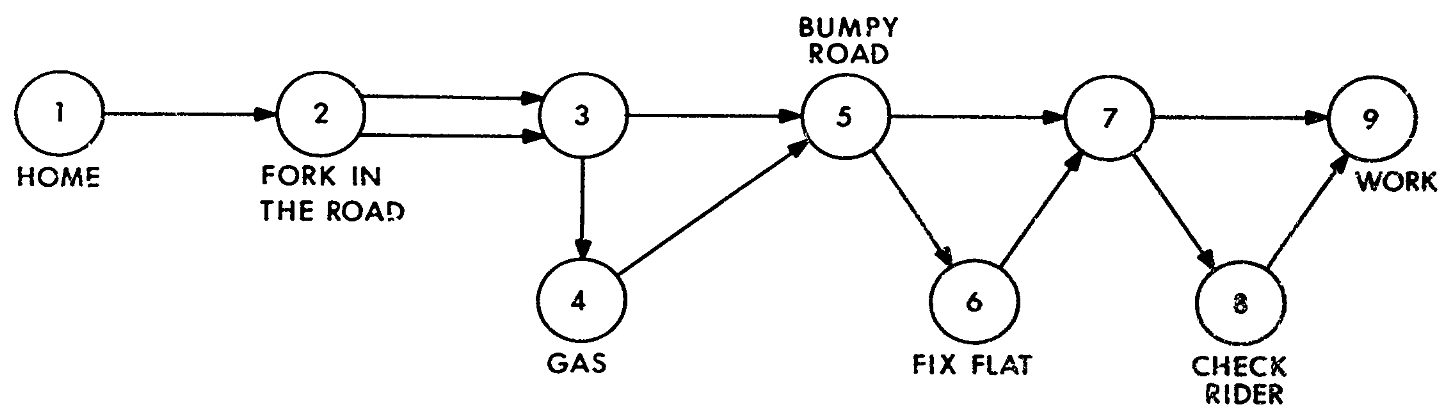


Figure 4.13 Network Representation of Route to Work.

Now that the worker has identified the route and its characteristics, he must assign probabilities to those nodes involving chance and time values to the arcs involving time. No costs are directly involved in the driving to work aspect of the problem.

The worker's past experience wi elapsed time required to drive the different segments of the road enables him to specify a most likely, optimistic and pessimistic completion time for each arc. Three time estimates are used for each arc in the network with the following three exceptions: stopping for gas, fixing a flat tire, and stopping to check for the rider.

All of these stops require about the same time. Therefore, the time for completion is assumed to be constant. In Table 4.1, all of the activity times are summarized.

TABLE 4.1 ACTIVITY COMPLETION TIMES

Triangularly Distributed Arc Times

Activity	Time Distribution Arguments (Minutes) Optimistic	Arc Identification	Most Likely	Pessimistic
From Home to Fork in the Road	10.0	(1,2)	12.0	14.0
Fork in the road to gas station (primary route)	5.0	(2,3)	6.0	7.0
Fork in the road to gas ,tation (alternate route)	11.0	(2,3)	13.0	15.0
Gas Station to Bumpy Road	3.0	(3,5)	4.0	5.0
Bumpy Road to Rider's House	4.0	(5,7)	5.0	6.0
Rider's House to Work	8.0	(7,9)	9.0	10.0

Constant Arc Times

Activity	Arc Identification	Time Required (Minutes)
Stop for gas	(3,4)	5.0
Change flat tire	(5,6)	15.0
Stop to check rider	(7,8)	5.0

In addition, the worker's experience enables him to subjectively estimate the probability of the events occurring for the chance nodes. In Table 4.2, a summary of the chance nodes and the workers estimate of the probability of occurrence is presented.

TABLE 4.2 PROBABILITY NODES

Event	Probability of Occurrence
Traveling Primary Route upon Reaching Fork	.90
Traveling Alternate Route upon Reaching Fork	.10
Stopping for gas if Alternate Route Taken	.20
Having a Flat Tire Along Bumpy Road	.05
Stopping for Rider if Alternate Route Taken	.20
Stopping for Rider if Gas Stop is Made	.40
Stopping for Rider After Flat Tire	.60

At this point, the worker has all the information required to run a network analysis of his route to work using the RISCA program. Since our objective is only to indicate how the output of this program could be used in a risk analysis or as a tool for decision-making purposes, there will not be a detailed description of the computer output. Only the relevant output will be presented. Once again, the interested reader is referred to Reference 3 for details.

Since in this example the worker is not interested in the path to work, only the distribution of arrival times, weighted over all possible outcomes, will be examined. Figure 4.14 presents the cumulative frequency histogram of this arrival time. The vertical axis is time to completion, and the horizontal axis is the probability that the true time is less than or equal to the time on the vertical axis. For instance, the probability of arriving at work in less than 35 minutes is approximately 0.45.

Assuming the worker always leaves his house at 0705 every morning and work starts at 0745, the probability of his being late on any given day is approximately 0.15. This is obtained as follows:

a. The worker always leaves his house at 0705. The probability of his being late is the probability that it will take over 40 minutes to get to work.

---

[3]Loc. Cit.

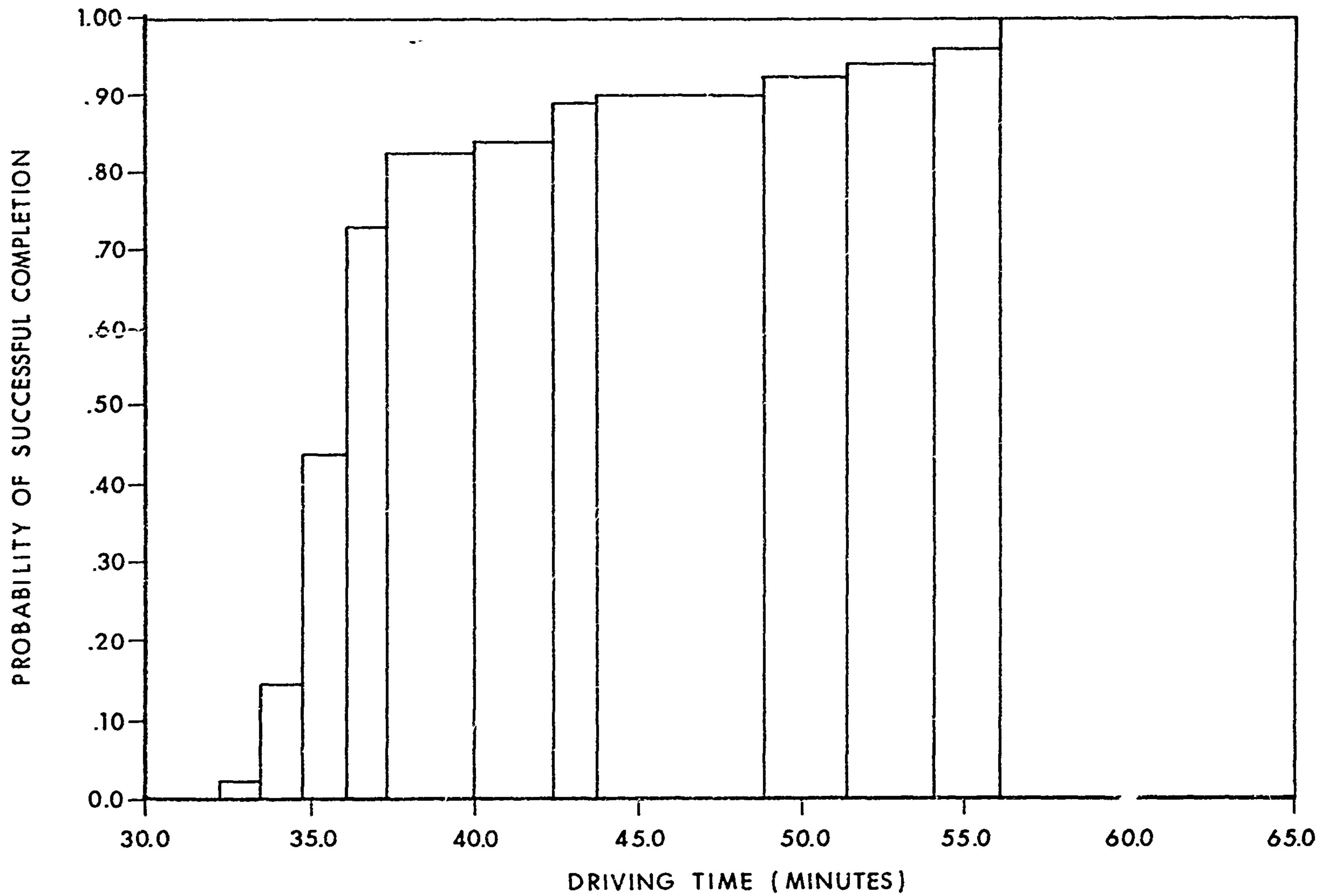


Figure 4.14. Cumulative Frequency Histogram of Completion Times for all Events.

b. Figure 4.14 gives cumulative probabilities like the probability of arriving at work in less than or equal to 40 minutes.

c. The probability that the worker is late is one minus the probability obtained in b.

If the probability of being late on any given day that the worker departs at 0705 is 0.15 then the risk of being late over 10 percent of the next 500 working days is the probability of being late 51 or more times. Let ℓ be the random variable representing the number of times late and p the probability of being late on a given day, then

$$P[51 \leq \ell \leq 500] = \sum_{\ell=51}^{500} \binom{500}{\ell} p^{\ell}(1-p)^{500-\ell}$$

is the risk of being late over 10 percent of the time. If the worker leaves at 0705, the above calculation results in a risk of 0.99.

Using this method it is possible to derive a risk profile for the worker as a function of departure time (See Figure 4.15). What should the worker do? The answer is not clear cut. It depends upon the worker's assessment of the value of additional leisure time versus the potential financial impact of being docked. However, given this risk profile the worker is in a position to consider the trade-offs rationally. For example, if the worker cannot possibly afford to be docked then the worker can select a departure time from the risk profile where the risk is approximately zero (i.e., leave 0653 or earlier).

Thus, this type of network analysis can be used to estimate risk and may be a valuable tool for making decisions.

4.4 SUMMARY

PERT provides a management tool that can help structure and discipline the R&D management process. In particular, PERT provides a method for identifying potentially high risk situations, and hence it can be a valuable tool for planning and scheduling uncertain program activities to meet program schedule objectives. In addition, PERT can provide guidance for the manager in controlling program activities by focusing his attention on critical activities. However, many of the assumptions that are generally made in applying PERT should be considered carefully before the technique is used.

Assuming that the events are deterministic may be fairly reasonable for short range planning and scheduling in a R&D program, but this assumption may break down for long range planning and scheduling. If the assumption does not apply, we may be removing a major element of the uncertainty. Not only will the network representation be unrealistic,

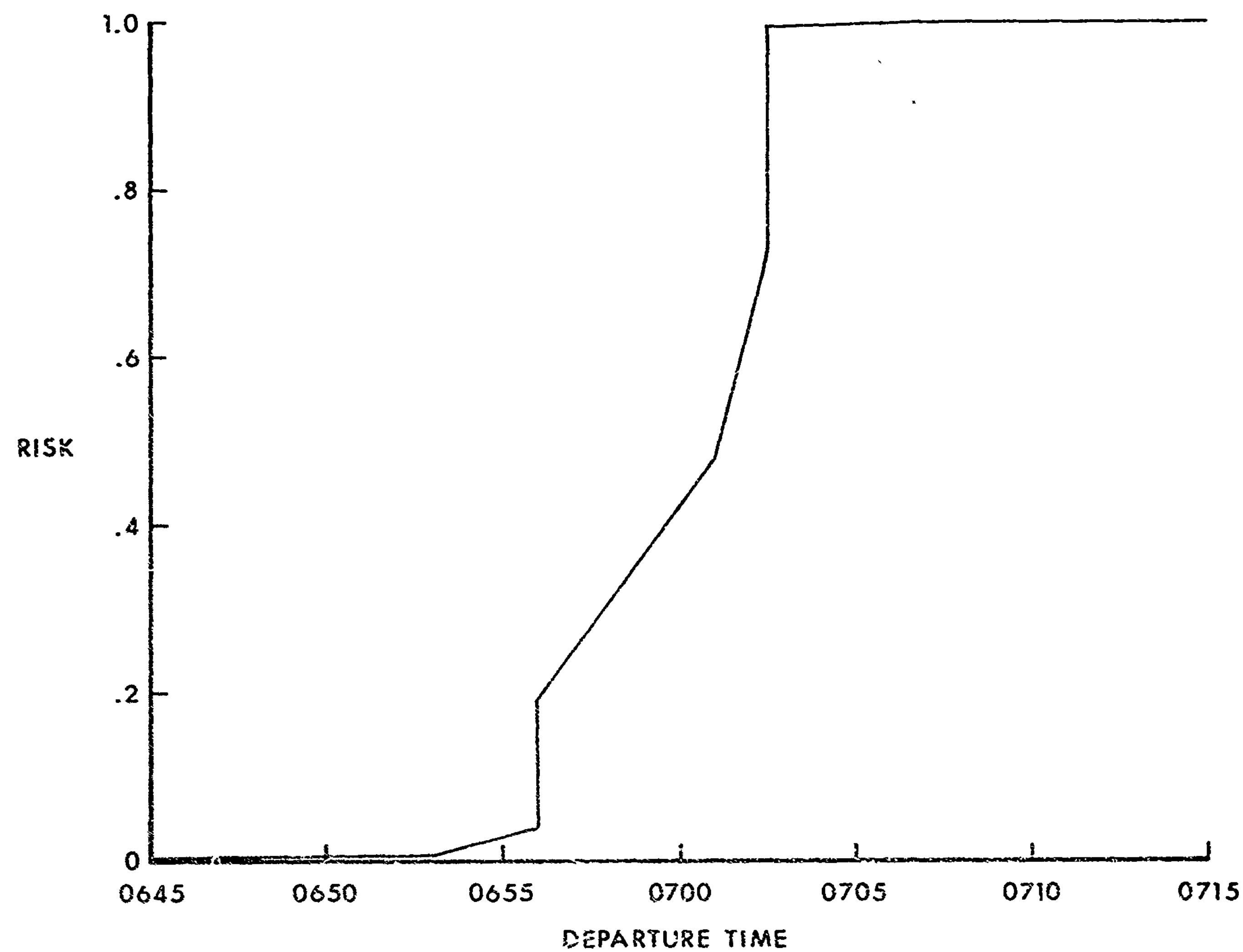


Figure 4.15 Risk Profile

but analyzing such a network may be fruitless and using the results may be dangerous. In addition, the assumption that the earliest completion time always occurs on the critical path may be misleading. The net effect of this assumption might be that a critical activity was not identified and hence was not closely controlled. In this case, the program could be delayed because of delays in this unidentified activity.

These two assumptions represent the greatest potential problems for PERT users and should be examined closely. However, simulation offers an alternative to making the second assumption and VERT offers an alternative when both assumptions are in question. So one should not feel stymied if these assumptions are suspect.

On the other hand, RISCA provides a framework for realistically modeling uncertainty in network representations and analyzing time, cost, and event uncertainty. In the context of risk analysis, RISCA and the other programs mentioned provide a method of quantifying (i.e., consolidation activities) program time and cost risks, and in the context of decision risk analysis these programs can provide the foundation of the structure for the evaluation for decision-making purposes.

Next page is blank.

LITERATURE CITED

1. ALMC, A Course of Instruction in Risk Analysis, ALM-3315-H, United States Army Logistics Management Center, Fort Lee, Virginia.

2. Mathematica, MATHNET, Preliminary Edition, August 1970, Princeton, New Jersey.

3. Brooks, W., Foster, W., and Maruyama, R.; MATHNET and RISCA (Network Analyzer Programs), A Users Manual, to be published as an AMSAA Technical Report, 1972, Aberdeen Proving Ground, Maryland.

BIBLIOGRAPHY

1. Ackoff, R. L., and Sasieni, M. W.; Fundamentals of Operations Research, 1968, John Wiley & Sons, Inc., New York, 1968.

2. ALMC, A Course of Instruction in Risk Analysis, ALM-3315-H, United States Army Logistics Management Center, Fort Lee, Virginia.

3. Brooks, W. J., Foster, W. V., and Maruyama, R. T.; MATHNET and RISCA (Network Analyses Programs), A Users Manual, Technical Memorandum to be published in 1972, U.S. Army Materiel Systems Analysis Agency, Aberdeen Proving Ground, Maryland.

4. Clarke, R. W.; The PERT Model for the Distribution of an Activity Time, No. 3, 1962, Operations Research, 10, pp 405-406.

5. Grubbs, F. E.; Attempts to Validate Certain PERT Statistics or Picking on PERT, 1962, Operations Research, 10, pp 912-915.

6. MacCummon, K. R., An Analytical Study of the PERT Assumptions, RM-3408-PR, 1962, Rand Corporation.

7. Malcolm, D. G., Roseboom, J. H., Clark, C. E., and Fayar, W.; Application of a Technique for Research and Development Program Evaluation, 1959, Operations Research, 7, pp 646-669.

8. Mathematica, MATHNET, Preliminary Edition, August 1970, Princeton, New Jersey.

9. Moeller, G. L.; VERT-A Tool to Assess Risk, 1972, U.S. Army Management Engineering Training Agency.

10. Pritsker, A. A. B., and Happ, W. W.; GERT: Graphical Evaluation and Review Technique Part I Fundamentals, May 1956, Journal of Industrial Engineering.

CHAPTER 5

BAYESIAN STATISTICS

5.1 INTRODUCTION

"A Bayesian believes that any quantity whose value he does not know is (to him) a random variable. He believes that it is possible, at any time, to express his state of knowledge about such a random variable in the form of a probability density function. As additional experimental evidence becomes available, Bayes' Theorem is used to combine this evidence with the previous probability density function in order to obtain a new posterior probability density function representing his updated state of knowledge. The probability density function expressing the analyst's state of knowledge serves as a quantitative basis for any decisions he is required to make." (Reference 1).

To motivate the usefulness of the Bayesian philosophy, consider the following hypothetical decision problem. Suppose that the US Army is involved in the early stages of a development program for a major weapon system. In the near future the decision must be made as to whether or not the Army will commit a sizeable portion of resources for the production buy of this system. This decision could involve millions of dollars over the system's life. One factor having a significant impact in such a decision is how well the system stacks up against a predetermined set of performance specifications. In this situation, there are generally two important questions facing the decision maker:

a. What are the chances that the system will meet the requirements specified? (i.e., What is the risk that the system will not meet the requirements?)

b. If the chances of the system meeting the requirement are small, what less stringent requirement could be met with an acceptable likelihood?

Unfortunately, testing of large systems is expensive both in terms of time and money. Consequently, initial production decisions are frequently made with very little production data. However, in most development programs there always exists other information that should be given consideration in these decisions. This may take the form of preproduction test data, engineering judgment, and experience with similar systems.

[1] Drake, Alvin W.; Bayesian Statistics for the Reliability Engineer, Proceedings of the National Symposium on Reliability and Quality Control, IEEE, 1966, pp 315-320.

In such instances, the Bayesian philosophy provides a rational framework for consolidating all this information (prior beliefs and current test data) into one simple concise statement concerning the chances of the system realizing the specified requirement.

Before examining the details of the Bayesisn procedure, consider the following few paragraphs taken directly from an article by F. J. Anscombe. They give a brief but concise historical development of both the Bayesian and the more orthodox or classical statistical philosophy.

> "During the last few years there has been a revival of interest among statistical theorists in a mode of argument going back to the Reverend Thomas Bayes[1] (1702-61), Presbyterian minister at Tunbridge Wells in England, who wrote an 'Essay towards solving a problem in the doctrine of chances,' which was published in 1763 after his death. Bayes's work was incorporated in a great development of probability theory by Laplace and many others, which had general currency right into the early years of this century. Since then there has been an enormous development of theoretical statistics, by R. A. Fisher, J. Neyman, E. S. Pearson, A. Wald and many others, in which the methods and concepts of inference used by Bayes and Laplace have been rejected.
>
> "The orthodox statistician, during the last twenty-five years or so, has sought to handle inference problems (problems of deciding what the figures mean and what ought to be done about them) with the utmost objectivity. He explains his favorite concepts, significance level, confidence coefficient, unbiased estimates, etc., in terms of what he calls probability, but his notion of probability bears little resemblance to what the man in the street means (rightly) by probability. He is not concerned with probable truth or plausibility, but he defines probability in terms of frequency of occurrence in repeated trials, as in a game of chance. He views his inference problems as matters of routine, and tries to devise procedures that will work well in the long run. Elements of personal judgment are as far as possible to be excluded from statistical calculations. Admittedly, a statistician has to be able to exercise judgment, but he should be discreet about it and at all costs keep it out of the theory. In fact, orthodox statisticians show a great diversity in their practice, and in the explanations they give for their practice; and so the above

remarks, and some of the following ones, are no better than crude generalizations. As such, they are, I believe, defensible. [Perhaps it should be explicitly said that Fisher, who contributed so much to the development of the orthodox school, nevertheless holds an unorthodox position not far removed from the Bayesian; and that some other orthodox statisticians, notably Wald, have made much use of formal Bayesian methods, to which no probabilistic significance is attached.]

"The revived interest in Bayesian inference starts with another posthumous essay, on 'Truth and Probability,' by F. P. Ramsey[2] (1903-30), who conceived of a theory of consistent behavior by a person faced with uncertainty. Extensive developments were made by B. de Finetti and (from a rather different point of view) by H. Jefferys. For mathematical statisticians the most thorough study of such a theory is that of L. J. Savage[3,4]. R. Schlaifer[5] has persuasively illustrated the new approach by reference to a variety of business and industrial problems. Anyone curious to obtain some insight into the Bayesian method, without mathematical hardship, cannot do better than browse in Schlaifer's book.

"The Bayesian statistician attempts to show how the evidence of observations should modify previously held beliefs in the formation of rational opinions, and how on the basis of such opinions and of value judgments a rational choice can be made between alternative available actions. For him probability really means probability. He is concerned with judgments in the face of uncertainty, and he tries to make the process of judgment as explicit and orderly as possible.

---

[1]Bayes, T., Essay Towards Solving a Problem in the Doctrine of Chances, reprinted with bibliographical note by G. A. Barnard, Biometrika, 45 (1958), 293-315,

[2]Ramsey, F. P., The Foundations of Mathematics, London: Rowtledge and Kegan Paul, 1931.

[3]Savage, L. J., The Foundations of Statistics, New York, John Wiley, 1954.

[4]Savage, L. J., Subjective Probability and Statistical Practice, to be published in a Methuen Monograph.

[5]Schlaifer, R., Probability and Statistics for Business Decisions: An Introduction to Managerial Economics Under Uncertainty, New York, McGraw-Hill, 1959." (Reference 2).

---

[2]Anscombe, F. J., American Statistician, Vol 15, 1961, pp 21-24.

Since Anscombe's article (Reference 2), Bayesian influence in statistical theory and practice has become more and more prevalent. Many works appear in the literature concerning the use of Bayesian thinking primarily in the areas of business decision making and reliability measurement. Of particular interest is a recent introductory text on Bayesian statistics by Samuel Schmitt (Reference 3). This text, written at a beginner's level, contains an excellent discussion of the Bayesian point of view. It also contains an extensive bibliography including a list of references by subject area along with an indication of the degree of difficulty of the respective work. Schmitt's text is highly recommended for an introduction to Bayesian analysis.

In the next section, a mathematical description of the Bayesian updating procedure is presented. Its purpose is to familiarize the reader with the basic theory, terminology, and notation encountered with the technique. This is followed by a detailed practical example illustrating the mechanics involved in applying the theory of the updating procedure. The example, concerning the development program of a missile system, is particularly representative of the type problem encountered in risk analysis within the US Army today. A section then follows which contains a discussion of the significant advantages and disadvantages of the Bayesian approach along with recommendations concerning its applicability in risk analysis.

5.2 DESCRIPTION OF THE BAYESIAN UPDATING PROCEDURE

Consider the Bayesian analysis of p, an unknown parameter of a postulated probabilistic model of a physical system. Assume that the experimental outcomes with the system can be treated as the values of a random variable x, the characteristic of interest. Based on past experience and all other available information, the Bayesian approach begins with the specification of a prior probability density function $f_p(p)$, that is a probability density function reflecting the analyst's prior beliefs about the value of the parameter, p. The assumed model specifies the probability density function for the sample value of the characteristic x, given the value of the parameter, p. Since p is being regarded as another random variable, the p.d.f. for the sample value of x with parameter p is written as the conditional p.d.f.,

$f_{x|p}(x_o|p_o)$ = conditional p.d.f. for the sample value of characteristic x, given that the value of parameter p is equal to p_o.

[2] Ibid.

[3] Schmitt, Samuel A.; Measuring Uncertainty - An Introduction to Bayesian Statistics, Addison Wesley Publishing Co., Inc., Reading Massachusetts, 1969.

Each time an experimental value of characteristic x is obtained, the continuous form of Bayes theorem* is used to obtain a posterior probability density function $f_{p|x}(p_o|x_o)$ representing the analyst's new state of knowledge about the value of the parameter, p. This posterior probability density function serves as the basis for any present decisions and also as the prior distribution for any future experimentation with the physical system.

As an example of the Bayesian updating procedure consider the weapon system development program mentioned in the introduction. Suppose that the system in question is a missile system and that production missile reliability is the crucial unknown decision variable. Since missile system test flight data are scored as successes and failures, the unknown reliability can be thought of as the success proportion, p, of a Bernoulli process. Thus, the conditional distribution of k successes in n trials, given p, is binomial and its probability density function can be expressed as

$$f_{k|p}(k|p) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

where k = 0,1,...,n.

Further, suppose that the Bayesian analyst feels that he can quantitatively specify all his prior beliefs about the missile reliability (p) in the form of a prior probability density function. In this case, his prior knowledge may be based on pre-production flight test data and/or expert engineering judgment and/or past experience with similar systems, etc.

For problems where a Bernoulli success proportion is the crucial decision variable, several arguments can be made in defense of using a a beta distributional representation for the prior. First, the beta is of a form which lends itself quite readily to the distribution of a proportion. Its range is the unit interval, it is unimodal and can be skewed in either direction. Thus by judicious choice of parameters, the beta probability density can easily be put into a form which will satisfactorily reflect one's prior beliefs. The beta prior assumption has the additional advantage that it simplifies the mathematics involved in the update procedure. Any last minute test results can readily be used to update the posterior beta distribution by merely repeating the update procedure with the posterior beta distribution now

---

$$^* \quad f_{p|x}(p_o|x_o) = \frac{f_{x,p}(x_o,p_o)}{f_x(x_o)} = \frac{f_{x|p}(x_o|p_o)\ f_p(p_o)}{\int_{p_o} f_{x|p}(x_o|p_o)\ f_p(p_o)\ dp_o}$$

assuming the role of the prior beta distirbution. Further, each update of the distribution reduces the impact of the subjectivity inherent in the initial prior assumption.

If p is assumed to have a beta distribution (one of the most widely used priors for this class of problems) with parameters ℓ and $m-\ell$, the probability density of p is given by

$$f_p(p) = C(\ell,m)\ p^{\ell-1}(1-p)^{m-\ell-1}$$

for $0 < p < 1$ and where $C(\ell,m)$* is the normalizing constant.

Using Bayes theorem to update this prior distribution of p with the experimental results (i.e., k successes in n trials), it follows that

$$f_{p|k}(p|k) = \frac{\frac{n!}{k!(n-k)!}\ p^k(1-p)^{n-k}\ C(\ell,m)\ p^{\ell-1}(1-p)^{m-\ell-1}}{\int_0^1 \frac{n!}{k!(n-k)!}\ p^k(1-p)^{n-k}\ C(\ell,m)\ p^{\ell-1}(1-p)^{m-\ell-1}\ dp}$$

After the necessary integration and algebraic manipulation the above equation reduces to:

$$f_{p|k}(p|k) = C(k+\ell,\ n+m)\ p^{k+\ell-1}(1-p)^{n-k+m-\ell-1}$$

which is again a beta distribution with parameters $k+\ell$ and $(n+m)$ $(k+\ell)$.

This posterior probability density function which represents the analyst's state of knowledge now serves as a quantitative basis for any decisions he is required to make.

In certain instances, the analyst may want to base his initial prior probability density function on pre-production test flights (where the parameters ℓ and m represent the number of successes and the number trials respectively) and then weight this prior distribution by applying a weighting factor to reflect other information such as engineering judgment and/or previous experience with similar systems, etc. This can be accomplished by simply applying the weighting factor, w, to the

---

* $C(\ell,m) = \frac{\Gamma(m)}{\Gamma(\ell)\ \Gamma(m-\ell)}$ where $\Gamma(m) = (m-1)!$ for positive integers m.

parameters ℓ and m which results in a prior of the form

$$f_p(p) = C(w\ell, wm)\, p^{w\ell-1}(1-p)^{w(m-\ell)-1}$$

for $0 < p < 1$. The updating procedure is then applied to this weighted prior distribution as in the unweighted case.

In a recent report by Atzinger and Brooks (Reference 4), the authors discuss a systematic approach for analyzing one's state of knowledge, taking into account certain basic considerations for constructing a prior distribution using the weighting factor approach. This report specifically addresses the class of problems characterized by success-failure type data. The interested reader is referred to that report for a detailed description of the method. For the purposes of this discussion, however, it will be assumed that an appropriate prior weighting factor has been determined.

Before proceeding, one further point should be mentioned. Although the emphasis in this discussion has been placed on the application of the Bayesian procedure to success-failure type data (Bernoulli process), the updating technique is certainly not restricted to this class.

In the next section, a more detailed version of the previously considered missile system example is examined to illustrate the details of the update procedure and to indicate how the decision maker can use information provided by the posterior beta distribution.

5.3 EXAMPLE

Assume the US Army is developing a surface-to-air missile to provide forward air defense for the Field Army. The tactical production decision is to be made in about a year, and to date there have been test firings with Research and Development rounds (40 firings) and Industrial Prototype rounds (30 firings). In the near future the Initial Production Tests will begin and it is anticipated that by the decision date there will be 20 test firings with production missiles. One important consideration facing the decision maker is the chance that the system will not meet the production missile reliability (R_M) requirement. Unfortunately, there will be only a limited amount of production test flight data available by the decision date, and if only production missile test data

[4]Atzinger, E. M., and Brooks, W. J.; Comparison of Bayesian and Classical Analysis for a Class of Decision Problems, Technical Report No. 59, April 1972, U.S. Army Materiel Systems Analysis Agency, Aberdeen Proving Ground, Maryland.

is used to estimate R_M, then a great deal of information is being ignored. Therefore, the problem is how can the pre-production missile data be meaningfully combined with the production data for decision-making purposes. To further compound the problem, the contractor is claiming that the quality control program at the manufacturing plant has been improved, and as a consequence the reliability is significantly higher than what has been demonstrated to date by pre-production rounds. The contractor's past performance and the fact that no concrete procedure changes have been instituted at the manufacturing plant make one suspect the claim. This latter subjective infromation serves as a basis for selecting an appropriate prior weighting factor.

In order to obtain an initial estimate of the prior distribution of production missile reliability, the test firing results to date must be scored. No example of a scoring criterion is provided here because it is not thought to be germane to this example. One point should be made however. The objective in developing a scoring criterion should be to remove all possible biases. For instance, if, as a result of the pre-production flights, design problems were diagnosed and corrected these flights should not be counted as observations. The results of a hypothetical scoring of the Research and Development, Industrial Prototype and Production missiles are summarized in Table 5.1. Based on this scoring, there are 40 observations for the pre-production rounds and 20 observations for production rounds (no tests do not count as observations).

TABLE 5.1 MISSILE FLIGHT FIRING SUMMARY

Type of Missile	Successes	Failures	No Tests
Research and Development	15	10	15
Industrial Prototype	10	5	15
Production	16	4	0

Using the 40 observations from the pre-production firings, an initial estimate of the prior distribution is given by

$$f_{R_M}(R_M) = \frac{\Gamma(40)}{\Gamma(25)\,\Gamma(15)} R_M^{24} \quad (1-R_M)^{14} \quad 0 < R_M < 1.$$

The next issue is to use some method to arrive at a meaningful weight for this prior distribution. This is a complex problem and should not be treated superficially. All relevant subjective information and experience with similar systems should be taken into consideration. For the purpose of this illustration, however, let us suppose that a rationale such as that described by Atzinger and Brooks (Reference 4) was applied and resulted in a prior weight equal to one half.

The parameters of the weighted prior beta distribution for this example are then 12.5 and 7.5 respectively. These are based on 25 successes in 40 pre-production missile test firings. For computational ease, the parameters can be rounded off to 13 and 8 without any apprecia-ble impact on the final results.

Given the parameters of the prior distribution ($\ell=13$ and $m-\ell=8$) and the update distribution parameters ($k=16$ and $n-k=4$), the posterior probability density function of R_M is

$$f_{R_M|16}(R_M|16) = \frac{\Gamma(41)\ R_M^{28}\ (1-R_M)^{11}}{\Gamma(29)\ \Gamma(12)}\quad 0 < R_M < 1.$$

Figures 5.1a and 5.1b depict both the posterior probability density and the corresponding cumulative distribution function.

Faced with a decision concerning an uncertain parameter R_M, and given its posterior beta distribution, the decision maker has several options available to him. He can use the cumulative distribution function of the variable R_M directly to address the following question: What is the probability of meeting a specific requirement for R_M? For For example, in this case the variable of interest, R_M, is a missile system's reliability. If the requirement is for R_M to be at least 0.8, the cumulative distribution in Figure 5.1b notes that the estimate of the probability that $R_M \geq 0.8$ is approximately 0.09.

While this is not a favorable result, one or more of the following steps can be taken:

a. Some less stringent requirement could be evaluated (e.g., the probability that R_M is greater than or equal to 0.7).

b. The distribution could be examined to determine the lower limit.

c. One could examine the sensitivity of the prior weight. In this case, however, the sensitivity analysis should not be conducted

---

[4] Loc. Cit.

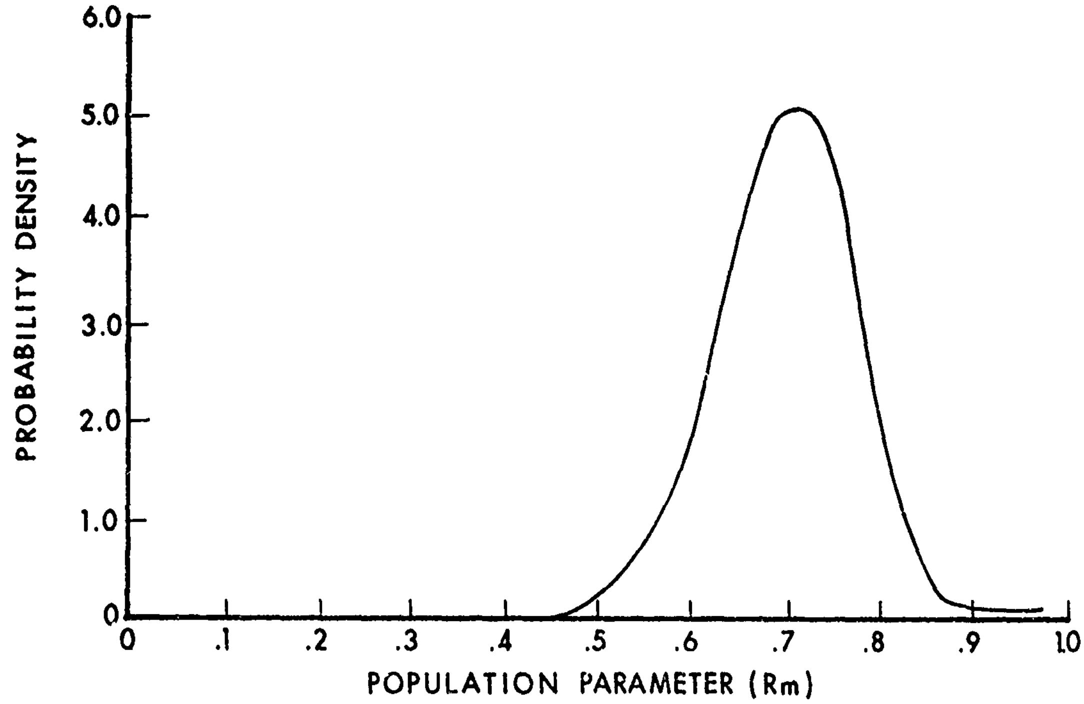


Figure 5.1a Beta Posterior Probability Density Function.

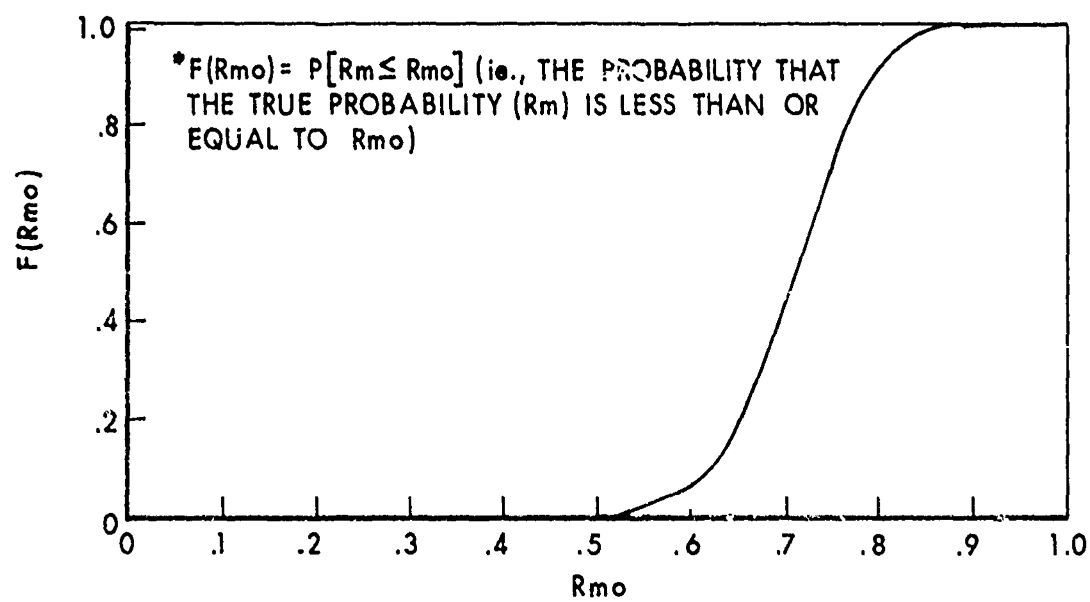


Figure 5.1b. Cumulative Posterior Beta Distribution.

indiscriminately (i.e., Don't play a numbers game). There should be legitimate reasons for changing the prior weight. These reasons will generally revolve around debate over the rationale (assumptions) for selecting the prior weight. For example, there may be two distinct opinions about the prior weight, one group may be optimistic (smaller weight) while the other group may be pessimistic (larger weight). After analyzing the rationale behind both of these opinions, the analyst may have selected a weight somewhere between these two schools of thought.

Suppose that in this case the prior weight is thought to be conservative. To examine the impact of being optimistic, a prior weight of one fourth is used. Using this weight, the parameters of the prior distribution are 6 and 4 respectively. This gives rise to the following posterior probability density function:

$$f_{R_M|16}(R_M|16) = \frac{\Gamma(30)\ R_M^{21}\ (1-R_M)^7}{\Gamma(22)\ \Gamma(8)} \quad 0 < R_M < 1.$$

Figures 5.2a and 5.2b depict this posterior probability density and the corresponding cumulative distribution function. Using the cumulative distribution function it follows that the probability that R_M is greater than or equal to 0.8 is now 0.2, a slight but insignificant increase.

Having examined this example of the updating procedure it is evident that the procedure is dependent on a significant amount of subjective analysis. This inherent subjectivity in the selection of the prior distribution and its corresponding prior weight is perhaps the basis for most of the attacks against the Bayesian philosophy. Non-Bayesians, who believe that the only legitimate types of probabilities are those that emanate from frequency-of-occurrence data, find it difficult to accept the idea of using subjective personalistic probabilities in forming a representative prior distribution. They contend that the selection procedure is rather arbitrary and that different analysts may come up with differing recommendations depending on their particular prior assumptions.

It is the Bayesian analyst's contention, however, that a decision maker faced with a real world decision will have intuition concerning the uncertain situation which is based on the externalities involved. He also feels that the most reasonable way for the decision maker to proceed is to heed his feelings and modify them on the basis of sample or experimental evidence. Certainly, he should not blind himself to a large portion of the information available merely on the basis that it may be subjective in nature, "If only objective probabilities have meaning then one cannot handle some of the most important uncertainties involved in problems of decision making." (See Reference 5).

[5] Hamburg, Morris; Statistical Analysis for Decision Making, Harcourt, Brace and World, Inc., New York, 1970.

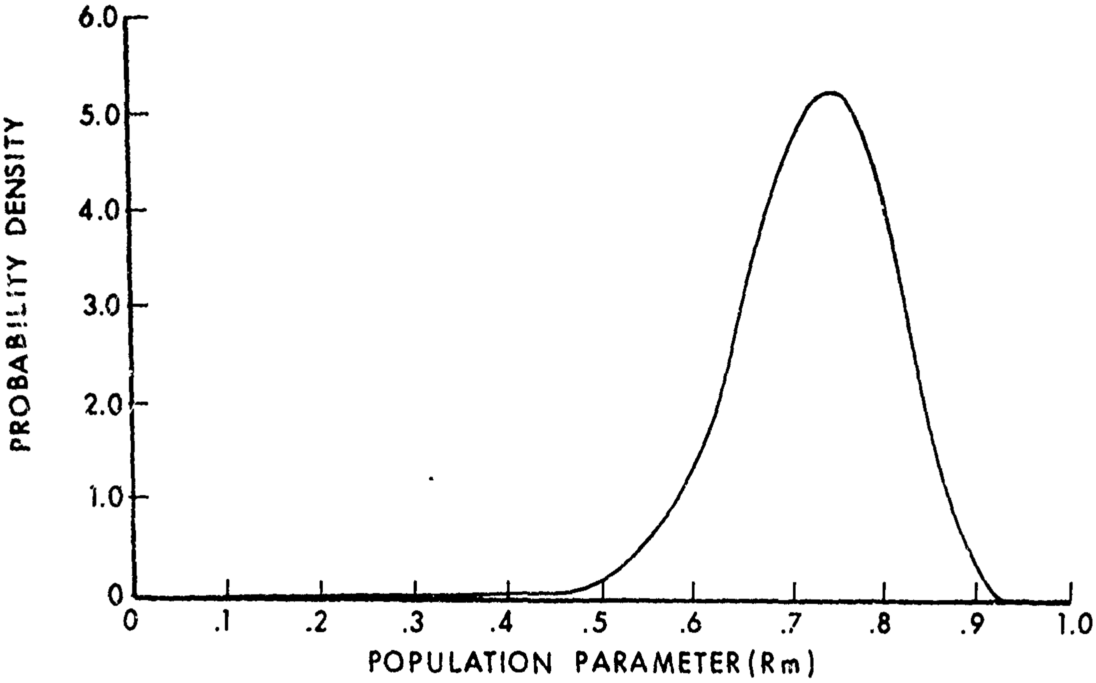


Figure 5.2a. Beta Posterior Probability Density Function.

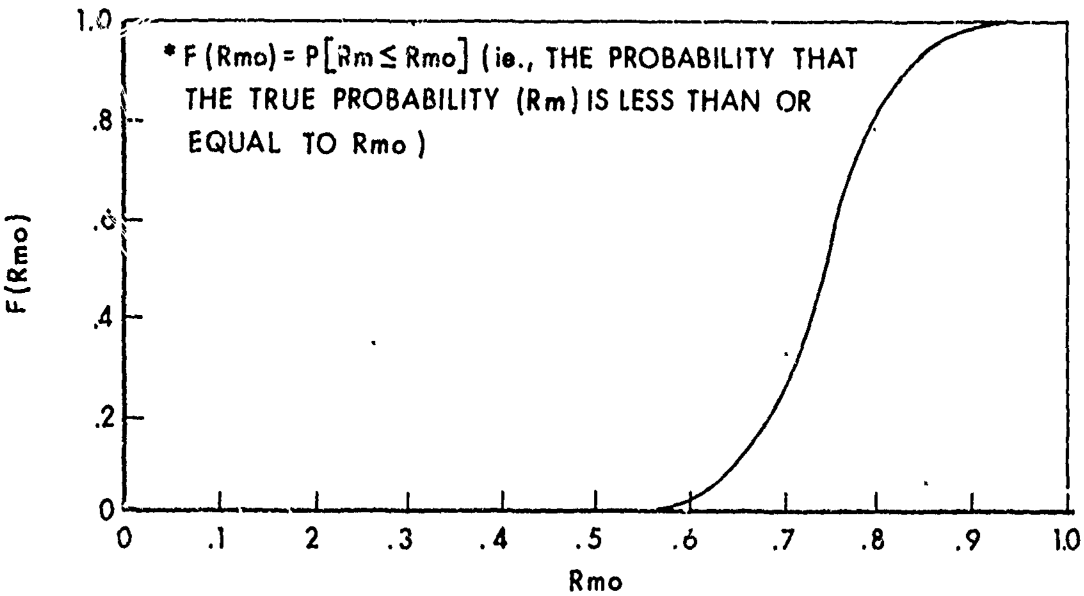


Figure 5.2b. Cumulative Posterior Beta Distribution.

The Bayesian agrees with the non-Bayesian that an inherent feature of the updating procedure is that different analysts may come up with different prior distributional assumptions. He feels, however, that this is a healthy situation rather than a weakness in the approach since the differences are clearly visible and can be used as a basis for further arbitration which may result in a more meaningful analysis.

For the production-buy decision in the missile system example just discussed, information such as that provided by pre-production test flight data, the contractor's quality control program, engineering judgment, and past experience with the contractor is perhaps some of the most crucial decision information available. Using the Bayesian procedure, the impact of this type information is reflected in the results thru the prior distributional assumptions.

In addition to the direct use of the posterior distribution of p for decision-making purposes, p may be a component of another variable of interest. In such an instance the posterior distribution can be used with standard statistical tools to determine the distribution of interest, or in cases where it is not practical to apply these tools, a Monte Carlo simulation can be used to examine the uncertainty in some function of the variable p. This function may or may not include elements of uncertainty other than p. Such a situation can be envisioned for the case previously considered where p represents a missile system's reliability.

Suppose for example the variable of interest is the single shot kill probability

$$P_{SSK} = R_{GSE} \cdot R_M \cdot P_{PF} \cdot M_L$$

where P_{SSK} = single shot kill probability for the missile system,

R_{GSE} = reliability of the ground support equipment,

R_M = reliability of the missile,

P_{PF} = probability of proper fuzing, and

M_L = probability of a kill given proper fuzing, a reliable missile, and reliable ground support equipment.

Since the uncertainty in the single shot kill probability is dependent on the uncertainty in the estimates of R_{GSE}, P_{PF} and M_L as well as the estimate of R_M, a Monte Carlo simulation may be used to examine this uncertainty. The uncertainty in the estimate of R_M can easily be introduced into a Monte Carlo simulation by sampling from the posterior cumulative distribution function of p.

5.4 SUMMARY AND CONCLUSIONS

In risk analysis, situations frequently exist where the analyst has available both objective test data and other relevant information based on the externalities of the problem. Often, due to cost and time constraints, there is only a limited amount of relevant test data available by the decision date. Thus, other factors such as previous test data, engineering judgment, experience with similar systems, etc., must be taken into consideration. In this context, Bayesian statistics provides the analyst with a tool for synthesizing (consolidation activities) this information into one probability distribution which can then be used directly to estimate the risks in question.

Unfortunately there seems to be some mystique that surrounds any application of Bayesian statistics. This is due in some instances to a disagreement with the Bayesian philosophy and in others to the lack of a true understanding of the mechanism of the Bayesian approach. Perhaps one of the most widely used arguments against the use of the Bayesian procedure is the apparent absence of a rational basis for constructing a prior distribution. For certain classes of problems, however, the argument has very little substance. Certainly in the missile system reliability problem, examined in detail in this article, there does exist a rational basis for selecting the prior distribution.

Thus the Bayesian approach, although not to be applied indiscriminately, should be given serious consideration as a viable tool for certain risk analysis and decision risk analysis applications.

LITERATURE CITED

1. Drake, Alvin W.; Bayesian Statistics for the Reliability Engineer, Proceedings of the National Symposium on Reliability and Quality Control, IEEE, 1966, pp 315-320.

2. Anscombe, F. J.; American Statistician, Vol 15, 1961, pp 21-24.

3. Schmitt, Samuel A.; Measuring Uncertainty - An Introduction to Bayesian Statistics, Addison Wesley Publishing Co., Inc., Reading, Massachusetts, 1969.

4. Atzinger, E. M., and Brooks, W. J.; Comparison of Bayesian and Classical Analysis for a Class of Decision Problems, Technical Report No. 59, April 1972, U.S. Army Materiel Systems Analysis Agency, Aberdeen Proving Ground, Maryland.

5. Hamburg, Morris; Statistical Analysis for Decision Making, Harcourt, Brace and World, Inc., New York, 1970.

Next page is blank.

END

DATE

FILMED

9-15-72